UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

André Saito Guerreiro

Perfomance analysis of latticed based MIMO receivers and FastICA in blind Massive MIMO Systems.

Análise de Desempenho de receptores baseados em reticulados para MIMO e FastICA em Sistemas MIMO cegos Massivos

Campinas

2018

UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

André Saito Guerreiro

Perfomance analysis of latticed based MIMO receivers and FastICA in blind Massive MIMO Systems.

Análise de Desempenho de receptores baseados em reticulados para MIMO e FastICA em Sistemas MIMO cegos Massivos

Thesis presented to the School of Electrical and Computer Engineering at the University of Campinas in partial fulfillment of the requirements for obtaining the Doctor degree in Electrical Engineering, in the area of Telecommunications and Telematics.

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica, na Área de Telecomunicações e Telemática.

Orientador: Prof. Dr. Gustavo Fraidenraich

Este exemplar corresponde à versão final da tese defendida pelo aluno André Saito Guerreiro, e orientada pelo Prof. Dr. Gustavo Fraidenraich

_____

Campinas

2018

Agência(s) de fomento e nº(s) de processo(s): CAPES, Não se aplica

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

G937p

Guerreiro, André Saito, 1986-
Perfomance analysis of latticed based MIMO receivers and FastICA in blind Massive MIMO systems / André Saito Guerreiro. – Campinas, SP : [s.n.], 2018.

Orientador: Gustavo Fraidenraich.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Sistemas MIMO. 2. Decodificadores (Eletrônica). 3. Estimação de sistemas. I. Fraidenraich, Gustavo, 1975-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Análise de desempenho de receptores baseados em reticulados para MIMO e fastICA em sistemas MIMO cego massivos.
Palavras-chave em inglês:
MIMO Systems
Decoders
System estimation
Área de concentração: Telecomunicações e Telemática
Titulação: Doutor em Engenharia Elétrica
Banca examinadora:
Gustavo Fraidenraich [Orientador]
Cláudio Ferreira Dias
Renato da Rocha Lopes
João Luiz Rebelatto
Walter da Cruz Freitas
Data de defesa: 20-12-2018
Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidato: André Saito Guerreiro RA: 072805

Data da Defesa: 20 de dezembro de 2018

Título da Tese: "Perfomance analysis of latticed based MIMO receivers and FastICA in blind Massive MIMO Systems." / "Análise de Desempenho de receptores baseados em reticulados para MIMO e *FastICA* em Sistemas MIMO cegos Massivos"

Prof. Dr. Gustavo Fraideinraich (Presidente, FEEC/UNICAMP)
Prof. Dr. Walter da Cruz Freitas Junior (GTEL/UFC)
Prof. Dr. João Luiz Rebelatto (UTFPR)
Prof. Dr. Renato da Rocha Lopes (Decom/Unicamp)
Dr. Claudio Ferreira Dias (Consultor)

Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

A minha família. O apoio e amor recebido me deram a segurança e paz para trilhar este caminho.

Acknowledgements

To all my family, for their love, support, and patience throughout my life.

To my advisor, Prof. Gustavo Fraidenraich for all the guidance and knowledge shared during the countless discussions.

To all my friends and colleagues, for making my days more enjoyable and easier to go through.

To all my teachers and professors, for providing me with the basic knowledge and skills that allowed me to present this work.

To CAPES, for the financial support.

Resumo

Estatísticas da taxa-soma dos decodificadores *Integer Forcing* (IF) e outros decodificadores baseados em reticulados para sistemas de Múltipla Entrada e Múltipla Saída (MIMO) são analisadas. Duas aproximações para a taxa-soma de decodificadores lineares são derivadas. A primeira aproximação é baseada no algoritimo Gauss-Lagrange para sistemas com duas antenas no transmissor e receptor (arranjo 2×2) e canais descorrelacionados. A segunda aproximação considera um sistema com um arranjo $n \times n$ de antenas, para o caso correlacionado e descorrelacionado e é baseado no segundo teorema de *Minkowiski*.

O desempenho de decodificadores IF e *Compute and Forward Transform* (CFT) são analisados na presença de erro de estimação de canal. Uma aproximação para a taxa-soma média na presença de erros de estimação de canal e canais com realização fixa é derivada. Uma aproximação para a taxa-soma ergódica dos decodificadores IF na presença de canais correlacionados e descorrelacionados também é derivada.

Decodificadores lineares IF atraíram atenção significativa devido ao seu potencial de atingir melhor desempenho do que outros decodificadores lineares, especialmente quando as matrizes de canal são aproximadamente singulares. No entanto, uma análise mais profunda de seu desempenho na presença de canais não determinísticos é necessária para que se possa quantificar sua vantagem em relação a decodificadores lineares clássicos e para que se possa corretamente projetar sistemas baseados nestes decodificadores.

Uma outra contribuição deste trabalho envolve decodificadores cegos em sistemas MIMO Massivos. Uma variação para o algoritmo *Fast Independent Component Analysis* (fastICA) que leva em consideração o formato das constelações para obter melhor desempenho é proposta.

Palavras-chaves: MIMO; MIMO Massivo; Decodificadores Lineares, Decodificadores *Integer Forcing*; Decodificador Cego; Erro de estimação do canal.

Abstract

The statistics of the sum-rate of Integer Forcing (IF) and other lattice-based Multiple Input Multiple Output (MIMO) systems are analyzed. Two approximations to the achievable sum-rate of the IF linear receiver and their respective analytical probability density functions (PDF) are derived. The first approximation is based on the Gauss-Lagrange algorithm for systems with two antennas at the transmitter and receiver (2×2 arrays) and uncorrelated channels. The second approximation considers an $n \times n$ array for both correlated and uncorrelated channels and its derivation is based on *Minkowiski*'s second theorem.

The performance of IF and Compute and Forward Transform (CFT) receivers is also analyzed under the presence of channel estimation errors. An approximation to their average sum-rate in the presence of these errors for fixed channel realizations is derived. An approximation to the Ergodic IF sum-rate for correlated and uncorrelated channels is also derived.

IF linear receiver has attracted significant attention recently due to their potential to perform better than other linear receivers, especially in the presence of channel matrices that are close to singular. However, a more in-depth analysis of its performance in the presence of non-deterministic channels is necessary in order to quantify its advantage over classical linear receivers and to correctly design systems that rely on these decoders.

Another contribution of this work involves blind decoding in Massive MIMO systems. We propose a variation to the fast Independent Component Analysis (fastICA) which takes into consideration the shape of the constellations to obtain better performance.

Keywords: MIMO; Massive MIMO; Linear Receivers; Integer Forcing Receivers; Blind Decoding; Channel Estimation Error.

“A man’s errors are his portals of discovery.”

James Joyce

List of Figures

Figure 2.1 – Example of a lattice. Here $\mathbf{g}_1 = \left(\sqrt{3}, 1\right)$ and $\mathbf{g}_2 = \left(-1, \sqrt{3}\right)$ 25

Figure 2.2 – Example of two different representations of the same lattice. The doted lines present the basis $\mathbf{v}_1 = \left(-1+2\sqrt{3}, 2+\sqrt{3}\right)$ and $\mathbf{v}_2 = \left(\sqrt{3}-1, -1+\sqrt{3}\right)$. The solid lines $\mathbf{g}_1 = \left(\sqrt{3}, 1\right)$ and $\mathbf{g}_2 = \left(-1, \sqrt{3}\right)$. It is clear that both basis span the same lattice, however $(\mathbf{g}_1, \mathbf{g}_2)$ has smaller vectors and is closer to orthogonality. 26

Figure 2.3 – Voronoi partition of the lattice spanned from the basis $\mathbf{g}_1 = \left(\sqrt{3}, 1\right)$ and $\mathbf{g}_2 = \left(-1, \sqrt{3}\right)$. 27

Figure 2.4 – Successive minima of the lattice $\mathbf{T} = (1.5, 0.7)\,; (1.8, 0, 6)$. We have $\|\mathbf{r}_1\| = \lambda_1$, $\|\mathbf{r}_2\| = \lambda_2$. 28

Figure 2.5 – Operation of conventional MIMO linear receivers. 31

Figure 2.6 – Operation of IF linear receivers. 32

Figure 3.1 – Frequency of channel realizations in which the output of one iteration of Gauss-Lagrange is optimal versus Total SNR (dB) 46

Figure 3.2 – Coefficient of Variation vs Number of antennas 47

Figure 3.3 – Ergodic Sum Rate versus Total SNR (dB) for 2×2 array with uncorrelated channels. 48

Figure 3.4 – Ergodic Sum Rate versus Total SNR (dB) for the 2×2 and 4×4 array with uncorrelated channels. 48

Figure 3.5 – Ergodic Sum Rate versus Total SNR (dB) for the 40×40 array. Black curves and blue curves are for uncorrelated and correlated channels respectively. 49

Figure 3.6 – Ergodic Sum Rate versus Total SNR (dB) for the 2×2 array and 4×4 with correlated channels. 49

Figure 3.7 – Outage Probability versus Sum Rate for the 2×2 complex array with uncorrelated channels at $s = 7dB$. 50

Figure 3.8 – Zoom of the low outage probability region versus Sum Rate for the 2×2 complex array with uncorrelated channels at $s = 7dB$. 51

Figure 3.9 – Outage Sum Rate 5% versus Total SNR (dB) for 2×2 and 4×4 arrays with uncorrelated complex channels 51

Figure 3.10–Outage Sum Rate 5% versus Total SNR (dB) for the 40×40 array. Black curves and blue curves are for uncorrelated and correlated channels respectively. 52

Figure 3.11–Outage Sum Rate 5% versus Total SNR (dB) for 2×2 and 4×4 arrays with correlated complex channels. 53

Figure 4.1 – Lattice points for $\Lambda((\mathbf{H}')^{\dagger})$ and $\Lambda(\mathbf{G})$. The solid arrows represent the vector basis before lattice reduction and the dotted arrows present the reduced vector basis. 58
Figure 4.2 – Number of antennas in each array vs Maximum correlation coefficient between elements of $\mathbf{\Omega}$ and $\mathbf{A}_e$ 62
Figure 4.3 – SNR vs Mean achievable IF sum rate $\left(\bar{R}_{IF_e}\right)$ for channel matrix $\mathbf{H}_s$. 63
Figure 4.4 – SNR vs Mean achievable CFT sum rate $\left(\bar{R}_{CFT_e}\right)$ for channel matrix $\mathbf{H}_s$. . . 64
Figure 4.5 – Signal to noise ratio versus Ergodic achievable IF sum rate for uncorrelated channels 65
Figure 4.6 – Signal to noise ratio versus Ergodic achievable IF sum rate for uncorrelated channels. Black curves are for $n = 4$, red for $n = 3$ and blue for $n = 2$ in a $n \times n$ array. The quantity $e = 0.1$ is used for all curves. 65
Figure 4.7 – Signal to noise ratio versus Ergodic achievable IF sum rate for correlated channels 66
Figure 4.8 – Signal to noise ratio versus Ergodic achievable IF sum rate for correlated channels. Black curves are for $n = 4$, red for $n = 3$ and blue for $n = 2$ in a $n \times n$ array. The quantity $e = 0.1$ is used for all curves 66
Figure 4.9 – Signal to noise ratio versus Ergodic achievable IF sum rate for uncorrelated channels with variable e. Parameters $\alpha = 0.2$, $\beta = 0.2$ and $L = n$. . 67
Figure 5.1 – SNR vs Symbol error rate. Scenario 1 (black curves) $L = 1, \tau = 7$. Scenario 2 (red curves) $L = 2, \tau = 4$ 75
Figure 5.2 – SNR vs Average number of rejected demixing vectors. Scenario 1 (red curves) $L = 1, \tau = 7$. Scenario 2 (red curves) $L = 2, \tau = 4$. 76

List of Tables

Table 2.1 – γ_ν values 29
Table 2.2 – Equalization matrix for the ZF and MMSE linear receivers 31

List of abbreviations and acronyms

BS *Base Station*

CDF *Cumulative Distribution Function*

CFT *Compute and Forward Transform*

CSI *Channel State Information*

DMT *Diversity Multiplexing Tradeoff*

HKZ *Hermite-Korkine-Zolotareff*

ICA *Independent Component Analysis*

ICA *Successive Interference Cancellation*

IF *Integer Forcing*

LLL *Lenstra–Lenstra–Lovász*

MIMO *Multiple Input Multiple Output*

ML *Maximum Likelihood*

MMSE *Minimum Mean Square Error*

PCICA *Phase Corrected Independent Component Analysis*

PDF *Probability Density Function*

SAGE *Space Alternating Generalized Expectation*

SIF *Successive Integer Forcing*

SIVP *Shortest Vector Problem*

SMP *Successive Minima Problem*

SNR *Signal to Noise Ratio*

ZF *Zero Forcing*

List of Symbols and Operators

$\mathbf{l}, \mathbf{L}$	Vectors are represented as boldface letters, matrices as upper-case boldface letters
$\mathbf{I}_n$	Identity matrix with n rows and columns
$\mathbf{0}_n$	Matrix with all zero elements with n rows and columns
Tr(.)	Trace operation
$(.)^{-1}$	Inverse of square a matrix
$(.)^{\dagger}$	Pseudo-inverse of a matrix
$(.)'$	Hermitian of a matrix
$\|\|.\|\|$	Norm operation
$\|.\|$	Absolute value operation
$\max(a, b)$	Largest value between real numbers a and b
$\log^{+}(.)$	Operation $\max(0, \log(.))$
$H(.)$	Heaviside step function
erf(.)	Error function
erfc(.)	Complementary error function
$\text{erfc}^{-1}(.)$	Inverse complementary error function
tahn(.)	Hyperbolic tangent function

$\lfloor . \rfloor$	Floor operation
$\lfloor . \rceil$	Round operation
$\Gamma(\cdot)$	The gamma function
j	Imaginary particle, $j = \sqrt{-1}$
$\mathbb{Z}$	Field of integer numbers
$\mathbb{Z}_p$	Field integer numbers modulo p
$\mathbb{C}$	Field of complex numbers
$\mathbb{E}[\cdot]$	Expectation operator

Contents

1 Introduction 20
1.1 Summary of Contributions 21
1.2 Related Work 22
1.3 List of Publications 23
2 Preliminaries 24
2.1 Lattices 24
2.1.1 Lattice Reduction 24
2.1.2 Voronoi partition 25
2.1.3 Successive Minima and Shortest Independent Vector Problem 26
2.2 Channel and System Model 29
2.2.1 Conventional Linear Receivers 30
2.2.2 Integer-Forcing Linear Receivers 32
2.2.2.1 Compute and Forward Transform (CFT) 34
2.2.2.2 Successive IF 35
3 Approximate Sum Rate for Integer-Forcing Receiver 36
3.1 Introduction 36
3.2 Proposed Approximations 37
3.2.1 IF and the Minkowski Minima 37
3.2.2 The Approximation for 2×2 array 37
3.2.2.1 Analysis of the Gauss-Lagrange Algorithm 38
3.2.3 The Approximation for $n \times n$ Array 39
3.2.3.1 Lower Bound to Compute and Forward Transform 40
3.2.4 The Error of the $n \times n$ Array Approximation 40
3.2.5 Approximate Degrees of Freedom of the $n \times n$ Array Approximation . 41
3.3 Distribution of the Approximations 41
3.3.1 Distribution of the 2×2 Array Approximation 41
3.3.1.1 Real Case 42
3.3.1.2 Complex Case 42
3.3.1.3 PDF of $\hat{R}_{IF_2}$ 42
3.3.2 Distribution of the $n \times n$ Array Approximation 43
3.3.2.1 Successive Integer Forcing Distribution 44
3.3.3 Ergodic Capacity, Outage and Outage Rate of the Approximations . . 44
3.4 Simulations 45
3.4.1 Gauss-Lagrange and second Minkowski Minima. 45

3.4.2 Mean Root Squared Error per Antenna . . . 46
3.4.3 Ergodic Rate $C_{\hat{R}_{IF}}$ . . . 46
3.4.3.1 Uncorrelated Channels . . . 47
3.4.3.2 Correlated Channels . . . 47
3.4.4 Outage Probability . . . 50
3.4.5 Outage Sum Rate . . . 50
3.4.5.1 Uncorrelated channels . . . 50
3.4.6 Correlated channels . . . 52
3.5 Chapter conclusions . . . 52
4 Approximate Sum Rate for Integer-Forcing Receiver with Imperfect Channel Estimation . . . 54
4.1 Introduction . . . 54
4.1.1 Effects of Imperfect CSI . . . 54
4.2 Approximation to the Average Achievable sum rate . . . 55
4.2.1 Average Achievable sum rate of IF and CFT receivers . . . 55
4.2.1.1 Relation of $\mathbf{A}_e$ and Ω . . . 56
4.2.1.2 Average achievable sum rate approximation . . . 57
4.2.2 Distribution of the Approximation of the Achievable IF sum rate under Imperfect CSI . . . 58
4.2.2.1 Uncorrelated channels . . . 59
4.2.2.2 Correlated channels . . . 60
4.2.2.3 Distribution of $R_{IF_{ap}}$ . . . 61
4.2.3 Ergodic sum rate of IF receivers with imperfect CSI . . . 61
4.3 Simulations . . . 62
4.3.1 $\mathbf{A}_e$ and Ω dependency . . . 62
4.3.2 Average achievable IF and CFT sum rate . . . 63
4.3.3 Ergodic IF achievable sum rate . . . 64
4.3.4 Variable e . . . 64
4.4 Chapter Conclusion . . . 67
5 Phase Corrected ICA for Uplink Massive MIMO . . . 68
5.1 Introduction . . . 68
5.2 System Model . . . 68
5.3 FastICA for blind communication systems . . . 69
5.3.1 PCA Whitening . . . 69
5.3.2 FastICA . . . 70
5.4 Phase Corrected ICA (PCICA) . . . 71
5.4.1 FastICA and Constellation Rotation . . . 71
5.4.2 Mean Value Criterion . . . 71

5.4.3 Performing the Z-Test 72
5.4.4 Phase Corrected ICA 74
5.5 Simulations 75
5.6 Chapter Conclusion 76
Conclusão 77

References 79
Appendices 84
A Appendix 85
A.1 Proof of Theorem 3.1 85
A.2 Proof of theorem 3.2 85
A.3 Proof of Theorem 3.3 87

Chapter 1

Introduction

The use of wireless communication has become commonplace in our lives. The flexibility of installation and use that arises from the lack of physical elements interconnecting communicating devices has led to the popularization of many applications which rely on it. In recent years, wireless data traffic has increased exponentially due to the increasing number of devices such as smartphones, laptops, tablets, and others. Mobile data traffic worldwide is expected to attain 49 exabytes per month by 2021 which is close to 2.9 times our current values [1]. However, this increase brings many challenges to future wireless systems, one of the major being the limited availability of frequency spectrum. Among several techniques studied to remedy this issue, multiple input multiple output (MIMO) is regarded as one of the most promising. By exploring the inherent spatial diversity of the MIMO channel it is possible to reliably multiplex different data streams over the same frequencies. This significantly increases the overall wireless channel capacity, as shown by the seminar papers by Foschini [2] and Telatar [3].

One of the key problems with MIMO systems lies on the receiver architecture. Joint maximum likelihood (ML) receivers are known to be optimal. However its implementation becomes prohibitively complex as the number of antennas increases and capacity-achieving channel codes are used. Linear receivers such as zero-forcing (ZF) or minimum mean square error (MMSE) receivers are widely used due to their reduced complexity, even though their performance is highly suboptimal when channel matrices are close to singular.

In [4] the authors introduced a method of using interference to obtain higher transmission rates on networks called Compute-and-Forward. Based on it, Integer-Forcing (IF) linear receivers have been introduced in [5], [6] and [7]. In the IF framework, each transmitter encodes an independent data stream and all transmitters are required to use identical lattice codes. This framework could correspond to an uplink scenario in a cellular transmission system, in which the transmitters are independent and send simultaneous data streams at the same frequency. It was shown that IF linear receivers perform better than other linear receivers, especially as channel matrices become close to singular. Moreover, using the results of [8] concerning the first successive minimum of a lattice induced by an i.i.d. Rayleigh channel matrix, it was proven that full diversity-multiplexing tradeoff (DMT) is attained.

Massive MIMO has also been a key research subject for next-generation wireless communications in recent years [9], [10]. Comprising a base station (BS) equipped with a large number of antennas which is several times greater than the number of active users. Its main appeal resides on the favorable propagation characteristic, which implies that different users channels are mutually orthogonal. In such conditions, simple linear precoding multiplexing is employed on the downlink and simple linear receivers on the uplink with almost no penalty to transmission rates.

These systems require channel state information (CSI) at the BS to achieve a high spectral efficiency. Usually, the system estimates the CSI making use of uplink training symbols. Users at the same cell send mutually orthogonal pilot sequences to ensure data stream separation at the BS. The accuracy of the estimation relates to the length of the pilot sequence but limits the number of users being served since the channel estimation process has to be repeated at every coherence period. Moreover, co-channel cells reusing the same pilot sequences in a multi-cell environment interfere in the channel estimation process. Therefore, the interference results in pilot contamination and reduces the performance of the system.

A great number of studies have explored the potential of blind and semi-blind decoding techniques to circumvent the issues regarding pilot based CSI estimation. In [11], the authors exploit the sparsity of the massive MIMO channel a blind decoding algorithm for systems with prior knowledge of channel and message distribution. In [12], the authors present a semi-blind algorithm that uses space-alternating generalized expectation maximization (SAGE) to decontaminate pilots. In [13], the authors propose a blind decoding method that uses an estimate of the aggregate out-of-cell channel covariance to remove interference.

Independent component analysis (ICA) based algorithms have gained significant attention among the several existing blind decoding algorithms. They are interesting due to the small amount of information necessary and the few restrictions on applications. It requires that the different sources of data be independent and not Gaussian distributed, and requires knowledge of the shape and average power of the constellation used by each user.

1.1 Summary of Contributions

Our work is divided into two parts. In Chapters 3, 4 we analyze the statistics of the sum-rate of IF and related receivers. In chapter 3, we derive two approximations to the achievable sum-rate of the IF linear receiver and their respective analytical probability

density functions (PDF) . The first approximation is based on the Gauss-Lagrange algorithm for systems with two antennas at the transmitter and receiver (2×2 arrays) and uncorrelated channels. The second approximation considers an $n \times n$ array for both correlated and uncorrelated channels and its derivation is based on *Minkowiski*'s second theorem. In order to validate our approximations, three metrics were evaluated, the mean value, the outage probability, and the outage rate.

In chapter 4, we analyze the effect of channel estimation error to the performance of IF and Compute and Forward Transform (CFT) receivers. We propose an approximation for the mean achievable sum rate of IF and CFT receivers with imperfect CSI. For IF receivers, we also derive the PDF of the approximation and present the expression of the ergodic IF achievable sum rate for uncorrelated and correlated channels. First, in the presence of complete CSI (chapter 3), and incomplete CSI (chapter 4) at the receiver.

In the second part of this work, we investigate blind decoding for Massive MIMO systems. In chapter 5, we propose a modified version of the fastICA algorithm called Phase Corrected ICA (PCICA). Exploring the format of commonly used constellations, it increases the performance of blind decoders based on ICA by iteratively selecting vectors that lead to estimated constellations that are more aligned with the axis.

1.2 Related Work

IF and other lattice based decoders have been extensively discussed in the specialized literature. Two important variations to the IF linear receivers are the CFT and Successive IF. CFT presented in [14] and [15] consists of adding an algebraic successive interference cancellation (SIC) method to the IF framework, increasing its attainable sum rate. Successive IF [16] uses noise predictor and is known to attain MIMO capacity with uniform power allocation. In [17] and [18], the performance of IF linear receivers was analyzed when Complex Lenstra–Lenstra–Lovász (LLL) , Hermite-Korkine-Zolotareff (HKZ) and Minkowski lattice basis reduction algorithms were used. In [19] an efficient algorithm to solve the successive minima problem (SMP) applied to IF linear Receivers was presented. In [20], the impact of channel variation on the IF linear receiver was studied and it was shown that it still presents an advantage over conventional linear receivers. In [21] and [22] precoders to the IF framework were introduced, and in [23] the uplink-downlink duality for Integer Forcing is discussed. The first precoder achieves the MIMO Capacity to within a constant gap, and the second precoder attains full diversity. Integer Forcing and Forward [24], a transceiver design for MIMO two way relay is introduced by the authors, with a variant suitable for channels with imperfect CSI.

Regarding receivers with imperfect CSI, in [25], the effects of channel estimation errors are analyzed and approximated bit error rates are derived for ZF receivers. In [26], the distribution of an approximation of the transmission rate for compute and forward systems with imperfect CSI is derived [24]. In [27], the effects of imperfect channel reciprocity and channel estimation errors for massive MIMO systems is analyzed.

As for ICA based blind decoders, in [28] the authors use a reduced number of pilots to remove the ambiguities inherent to ICA. In [29], statistics of the channel are used in order to remove permutation ambiguities in ICA. In [30], the authors use a modified ICA algorithm which reduces complexity with a small penalty to performance.

1.3 List of Publications

- A. S. Guerreiro , G. Fraidenraich, S. Kumar, "Approximate Sum Rate for Integer-Forcing Receiver," in IEEE Transactions on Communications, vol. 65, no. 11, pp. 4899 - 4910, July 2017.
- A. S. Guerreiro , G. Fraidenraich, L. L. Mendes, "Approximate Sum Rate for Integer-Forcing Receiver with Imperfect Channel Estimation", in IEEE Transactions on Vehicular Technology, vol. 67, no. 12, pp. 11767 - 11775, Dec 2018.
- A. S. Guerreiro , G. Fraidenraich, "Phase Corrected ICA for Uplink Massive MIMO" in IEEE Signal Processing Letters, vol. 25, no. 12, pp. 1810 - 1814, Dec 2018.

Chapter 2

Preliminaries

In this chapter, we review some concepts that are essential to the derivation of our work. We review basic lattice theory focusing on results that are important for the comprehension of IF receivers. We also introduce the channel and system model that are commonly used to describe MIMO linear receivers. The same model is used throughout chapters 3 and 4. Finally, we present the operation of linear receivers, also detailing the operation of IF linear receivers.

2.1 Lattices

A lattice is a discrete subgroup of $\mathbb{R}^n$. It is formed by a set of points which is closed under reflection and real addition. Therefore, if a point $\mathbf{p}_1$ is in the lattice, then so is its reflection $-\mathbf{p}_1$. And if two points $\mathbf{p}_1$ and $\mathbf{p}_2$ are in the lattice, so is their vector sum $\mathbf{p}_1 + \mathbf{p}_2$. These properties imply that all lattices also contain the origin, as the sum of $\mathbf{p}_1$ and $-\mathbf{p}_1$ (which are elements of the lattice due to reflection) is $\mathbf{0}$. It also imply that lattices are countably infinite sets, as they must contain all integer combinations of $a\mathbf{p}_1 + b\mathbf{p}_2$ for $a, b \in \mathbb{Z}$. Having established the basic properties of lattices, we can provide a more formal and constructive definition as,

Definition 2.1.1. A n-dimensional lattice Λ is defined by a set of n linearly independent set of column vectors (basis) $\mathbf{g}_1, \ldots, \mathbf{g}_n$ in $\mathbb{R}^k$. The lattice is constructed by the infinite set of points in $\mathbb{R}^k$ of the form $\{\mathbf{x} = \mathbf{G}\mathbf{u} \,|\, \mathbf{u} \in \mathbb{Z}^n\}$, where $\mathbf{G}$ is the generator matrix of Λ such that its columns are $\mathbf{g}_m$ $(m = 1, \ldots, n)$. We denote the lattice as $\Lambda(\mathbf{G})$.

Figure 2.1 presents an example of a lattice and shows the two vectors of the generator matrix.

2.1.1 Lattice Reduction

For a given lattice $\Lambda(\mathbf{G})$, any matrix $\bar{\mathbf{G}} = \mathbf{G}\mathbf{U}$ also generates the same lattice $(\Lambda(\mathbf{G}) = \Lambda(\bar{\mathbf{G}}))$ as long as $\mathbf{U}$ is a unimodular matrix. An unimodular matrix has integer elements and $|\det(\mathbf{U})| = 1$. Since the basis are not unique, it is possible to search for the

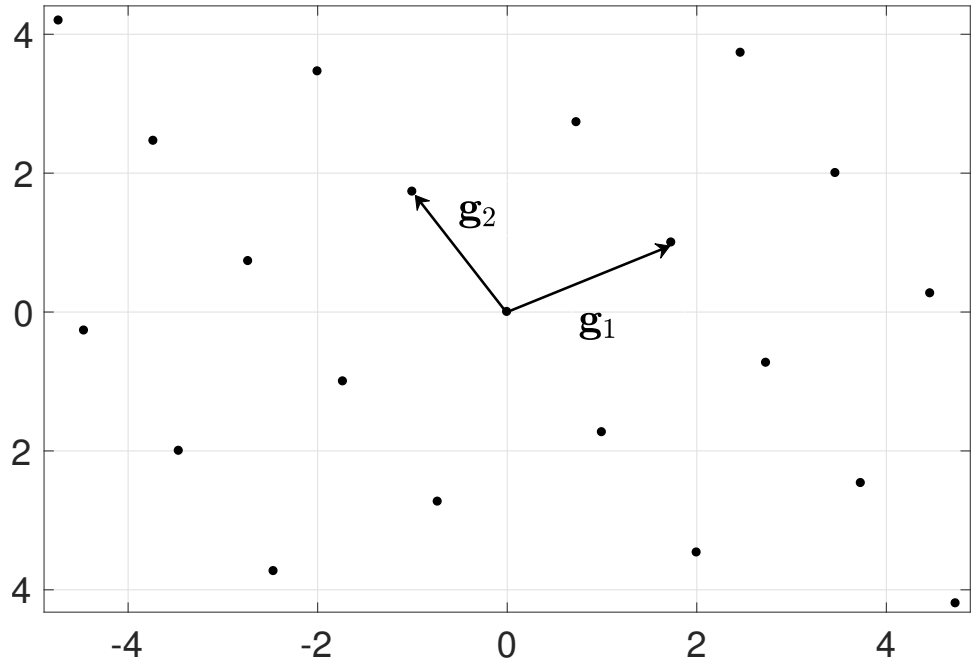

Figure 2.1 – Example of a lattice. Here $\mathbf{g}_1 = \left(\sqrt{3}, 1\right)$ and $\mathbf{g}_2 = \left(-1, \sqrt{3}\right)$

a generator matrix that is better in some sense. Although not a very precise definition, according to [31] a good rule of thumb for a good basis is,

- Basis vectors $(\bar{\mathbf{g}}_1, \ldots, \bar{\mathbf{g}}_n)$ are shortest possible,
- Basis vectors are nearly orthogonal.

The first criterion is related to numerical stability and the second is useful for reducing complexity in some problems, such as searching for the closest lattice point to a given point in space. This problem is closely linked to coding and decoding of lattice points. Figure 2.2 presents the same lattice of Figure 2.1 with two different basis. Its is clear that one representation has shorter and more orthogonal vectors. There is a great number of algorithms that attempt to find shorter basis that satisfy some near-orthogonality criterion. The LLL algorithm, HKZ, and Minkowski [32] algorithms are some of the most used for communications.

2.1.2 Voronoi partition

Using lattices, it is possible the divide the Euclidean space into congruent cells. Although many partitions are possible, one of the most used is the Voronoi partitions which uses nearest-neighboor rule. The distance from a point $\mathbf{x}$ to Λ is,

$$\|\mathbf{x} - \Lambda\| \triangleq \min_{\mathbf{p} \in \Lambda} \|\mathbf{x} - \mathbf{p}\| \tag{2.1}$$

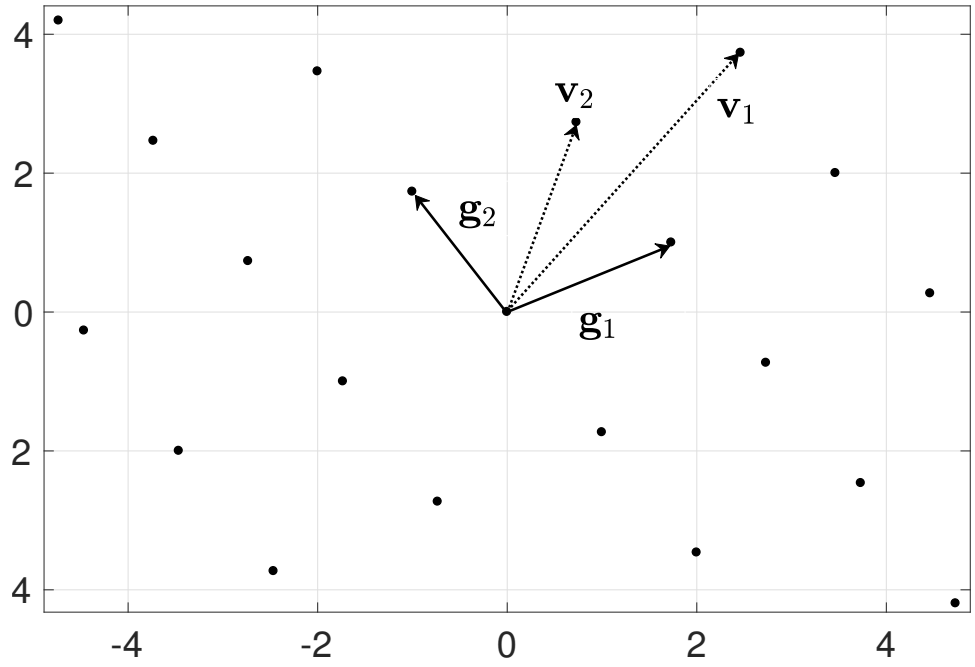

Figure 2.2 – Example of two different representations of the same lattice. The doted lines present the basis $\mathbf{v}_1 = \left(-1+2\sqrt{3}, 2+\sqrt{3}\right)$ and $\mathbf{v}_2 = \left(\sqrt{3}-1, -1+\sqrt{3}\right)$. The solid lines $\mathbf{g}_1 = \left(\sqrt{3}, 1\right)$ and $\mathbf{g}_2 = \left(-1, \sqrt{3}\right)$. It is clear that both basis span the same lattice, however $(\mathbf{g}_1, \mathbf{g}_2)$ has smaller vectors and is closer to orthogonality.

Here the norm operator $\|.\|$ calculates the Euclidean distance. The nearest-neighbor quantizer maps the point $\mathbf{x}$ to its closest lattice point is defined as,

$$Q_\Lambda^{NN}(\mathbf{x}) = \arg\min_{\mathbf{p}\in\Lambda} \|\mathbf{x}-\mathbf{p}\|. \tag{2.2}$$

Finally the Voronoi cell $\mathbb{V}_\mathbf{p}$ is the set of all points which are quantized to $\mathbf{p}$,

$$\mathbb{V}_\mathbf{p} = \{\mathbf{x} : Q_\Lambda^{NN}(\mathbf{x}) = \mathbf{p}\} \tag{2.3}$$

The volume of any Voronoi cell of $\Lambda(\mathbf{G})$ can be calculated as,

$$\text{Volume}(\mathbb{V}_\mathbf{p}) = \det(\Lambda) = \sqrt{\det(\mathbf{G}'\mathbf{G})}. \tag{2.4}$$

Figure 2.3 presents the Voronoi partition of the lattice presented in the previous figures.

2.1.3 Successive Minima and Shortest Independent Vector Problem

A parameter of great interest in lattice theory is the first Minkowski minima λ_1. It is defined as the smallest Euclidean distance r such that the lattice points inside a ball of radius r span a space of dimension 1. This definition is generalized, leading to the definition of successive minima.

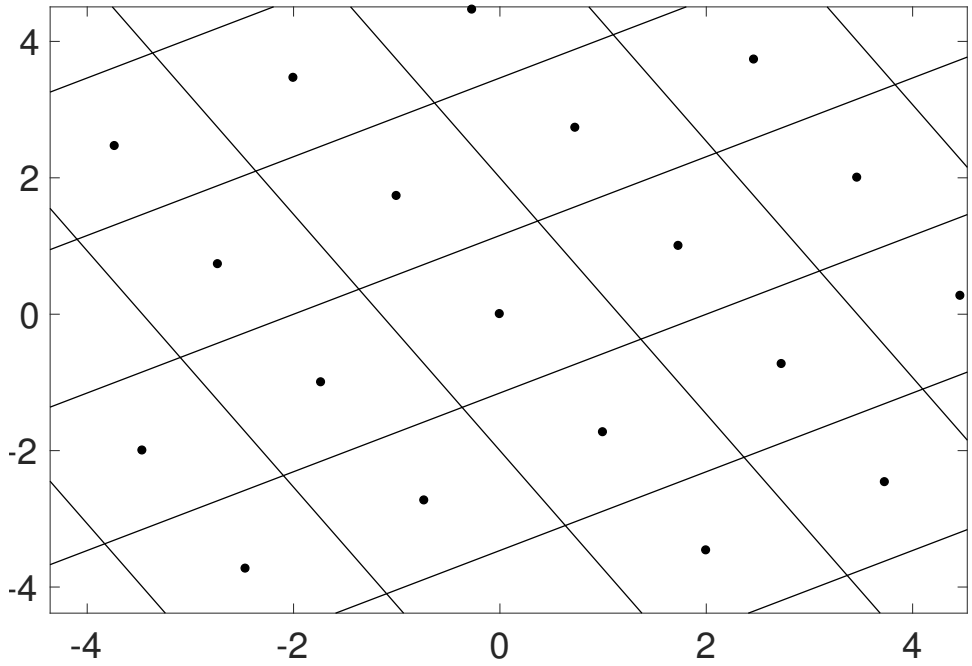

Figure 2.3 – Voronoi partition of the lattice spanned from the basis $\mathbf{g}_1 = \left(\sqrt{3}, 1\right)$ and $\mathbf{g}_2 = \left(-1, \sqrt{3}\right)$.

Definition 2.1.2. Given a lattice Λ in which its generator matrix has rank n, the i-th successive minima ($1 \leq 1 \geq n$) is defined as,

$$\lambda_i(\Lambda) = \inf\{r | \dim(\text{span}(\Lambda \cup \mathbb{B}(0, r))) \geq i\}. \tag{2.5}$$

Here $\mathbb{B}(0, r) = \{\mathbf{x} \in \mathbb{R}^k, \|\mathbf{x}\| \leq r\}$, which represents the closed ball of radius r around $\mathbf{0}$.

Given the same lattice Λ, the shortest independent vector problem (SIVP) consists on the search for the n linearly independent vectors $(\bar{\mathbf{g}}_1, \ldots, \bar{\mathbf{g}}_n)$ with the following constraint to its length,

$$\max_i \|\bar{\mathbf{g}}_i\| \leq \lambda_n(\Lambda). \tag{2.6}$$

Figure 2.4 presents an example of a lattice and its successive minima. For the two-dimensional real lattices, the Gauss-Lagrange algorithm is known to solve the SIVP. Algorithm 1 presents a slightly modified version of the Gauss-Lagrange algorithm that is also suitable for two-dimensional complex lattices. Overall the SIVP is considered to be NP-complete, therefore it is useful to present a bound to the successive minima using the second Minkowski theorem.

Definition 2.1.3. For any full-rank lattice Λ of rank n,

$$\sqrt{\gamma_n} \det(\Lambda)^{1/n} \geq \left(\prod_{i=1}^{n} \lambda_i(\Lambda) \right)^{\frac{1}{n}}. \tag{2.7}$$

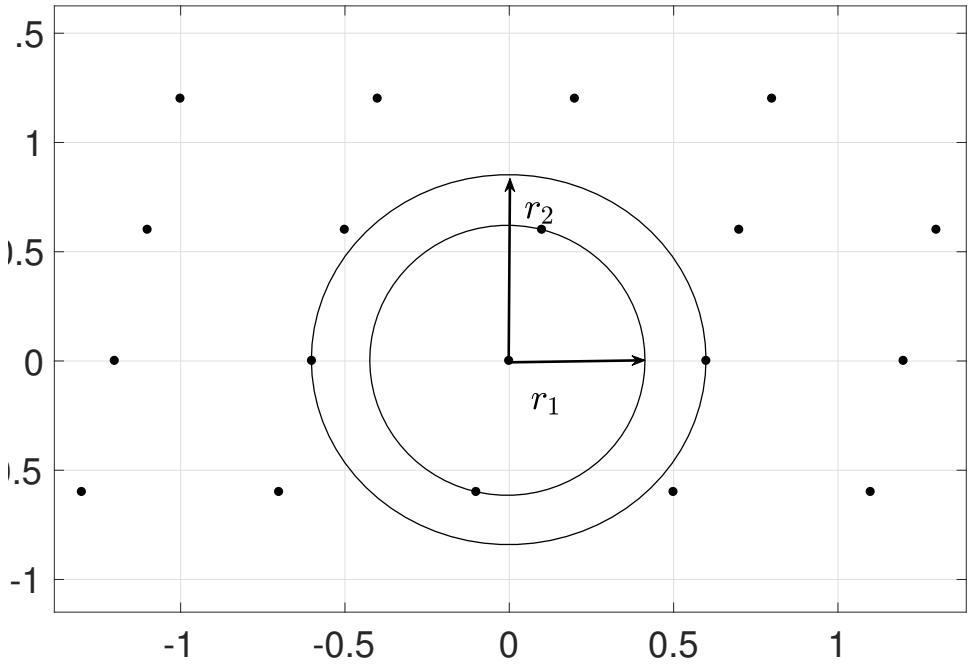

Figure 2.4 – Successive minima of the lattice $\mathbf{T} = (1.5, 0.7); (1.8, 0, 6)$. We have $\|\mathbf{r}_1\| = \lambda_1$, $\|\mathbf{r}_2\| = \lambda_2$.

The constant γ_ν is the Hermite constant in dimension ν is defined as,

$$\gamma_= \sup_{\Lambda} \left(\frac{\lambda_1(\Lambda)}{\det(\Lambda)^{1/\nu}} \right)^2, \tag{2.8}$$

where Λ ranges over all ν-dimensional lattices. Its values are known for $\nu \in [1, 8]$ and $\nu = 24$ [33]. Table 2.1 presents these values. There is also an important upper bound which

Data: Column vectors $(\mathbf{g}_1, \mathbf{g}_2) \in \mathbb{R}^2$ or $\mathbb{C}^2$ forming a base for lattice $\mathbf{G}$.
Result: Column vectors $(\mathbf{g}_1, \mathbf{g}_2)$ forming a base for $\bar{\mathbf{G}}$.
initialization;
if $||\mathbf{g}_2|| < ||\mathbf{g}_1||$ **then**
  Swap $\mathbf{g}_1$ and $\mathbf{g}_2$
end
$G_1 = ||\mathbf{g}_1||^2$;
$\mu = \left|\mathbf{g}_1'.\mathbf{g}_2\right| / G_1$;
$\mathbf{g}_2 = \mathbf{g}_2 - \lceil \mu \rfloor \mathbf{g}_1$;
$G_2 = ||\mathbf{g}_2||^2$;
while $G_2 < G_1$ **do**
  Swap $\mathbf{g}_1$ and $\mathbf{g}_2$;
  $G_1 = G_2$;
  $\mu = \left|\mathbf{g}_1'.\mathbf{g}_2\right| / G_1$;
  $\mathbf{g}_2 = \mathbf{g}_2 - \lceil \mu \rfloor \mathbf{g}_1$;
  $G_2 = ||\mathbf{g}_2||^2$;
end
Return $(\mathbf{g}_1, \mathbf{g}_2)$;

Algorithm 1: Gauss-Lagrange Algorithm

Table 2.1 – γ_v values

v	γ_v
2	$\sqrt{4/3}$
3	$2^{1/3}$
4	$4^{1/4}$
5	$8^{1/5}$
6	$(64/3)^{1/6}$
7	$64^{1/7}$
8	2
24	4

is asymptotic on v given by,

$$\gamma_v \leq \frac{1.744v}{2\pi e}. \tag{2.9}$$

2.2 Channel and System Model

We consider the existence of n independent data streams (or messages), $w_1, ..., w_{n_t}$ drawn from the same alphabet $\mathbb{Z}_p^k$, p being a prime number and k the message length. The m-th data stream w_m is mapped onto a length l channel input x_m by the m-th encoder. Uniform power allocation is considered across transmit antennas, i.e.,

$$\frac{1}{l}\|x_m\|^2 \leq s, \tag{2.10}$$

where s represents the average signal to noise ratio (SNR) per antenna considering the noise is normalized to unit variance. Also, each of the encoders transmits at the same rate

$$R_{TX} = \frac{k}{l}\log(p). \tag{2.11}$$

It is considered that the system has n_t transmitting and n_r receiving antennas. The $n_t \times l$ matrix $\mathbf{X}$ represents the transmitted signals where the m-th row of this matrix represents the signals from the m-th data stream. Channel gains are represented by the $n_r \times n_t$ matrix $\mathbf{H}$ and the additive white noise is represented by the $n_r \times l$ matrix $\mathbf{Z}$, its entries are considered to be i.i.d as $\mathscr{CN}(0,1)$. The received signal $\mathbf{Y}$ observed across the n_r receiving antennas is written as,

$$\mathbf{Y} = \mathbf{HX} + \mathbf{Z}, \tag{2.12}$$

It is important to note that the elements of $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ and $\mathbf{H}$ may be real or complex depending on whether the transmitter chooses to map its codewords into real or complex field.

We assume that the elements of $\mathbf{H}$ are i.i.d. Gaussian distributed (with Rayleigh distributed envelope for the complex case) and remains constant through the transmission block l. The knowledge of each of its realizations is only available at the receiver side. Under open loop operation, it is considered that the statistics of $\mathbf{H}$ are available at all transmitters. This is necessary to ensure that the transmitters use a transmission rate that leads to small outage probability. If a limited feedback is available, the receiver informs all transmitters the appropriate transmission rate, ensuring that sum-rate is attained [16].

In some sections, we consider two different scenarios concerning channel correlation. In the uncorrelated scenario, the entries of $\mathbf{H}$ are independent with zero mean and unit variance. When the correlation among the transmitting antennas exists, the rows of $\mathbf{H}$ are independent random vectors, but the elements of each row are correlated with zero mean and same covariance matrix $\mathbf{\Sigma}$. Each element of this covariance matrix can be expressed as,

$$\mathbf{\Sigma}_{(i,j)} = \mathbb{E}\Big[(\mathbf{h}_i - \mathbb{E}(\mathbf{h}_i))\Big(\mathbf{h}'_j - \mathbb{E}\Big(\mathbf{h}'_j\Big)\Big)\Big], \tag{2.13}$$

where $i \neq j$, $i = 1, \ldots, n_t$ and $\mathbf{h}_m$ represents the m-th column of $\mathbf{H}$. The correlation matrix ($\mathbf{\Sigma}$) is normalized and thus all its diagonal elements are unity.

Throughout this work, one of the main parameters of analysis is the achievable rate defined as,

Definition 2.2.1. (**Achievable Rate**) A sum-rate $R(\mathbf{H})$ is achievable if for any $\epsilon > 0$ and l large enough, there exists an encoder and decoder such that reliable communication decoding is possible,

$$\mathbb{P}((\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_n) \neq ((\mathbf{w}_1, \ldots, \mathbf{w}_n)) < \epsilon, \tag{2.14}$$

so long as the total rate does not exceed $R(\mathbf{H})$,

$$n_t * R_{TX} \leq R(\mathbf{H}). \tag{2.15}$$

2.2.1 Conventional Linear Receivers

Rather than jointly processing the observed signals from all the antennas through l channel usages, as the ML linear receiver, conventional linear receivers first attempt to decorrelate the transmitters data streams that are spatially coupled by the MIMO channel and then recover each individual data streams using parallel single input, single output (SISO) decoders. Upon receiving $\mathbf{Y}$, they apply an $n_r \times n_r$ equalization matrix $\mathbf{B}$,

$$\bar{\mathbf{Y}} = \mathbf{BY} = \mathbf{BHX} + \mathbf{BZ}. \tag{2.16}$$

Figure 2.5 illustrates the operation of these decoders. The aspect that differentiates ZF linear

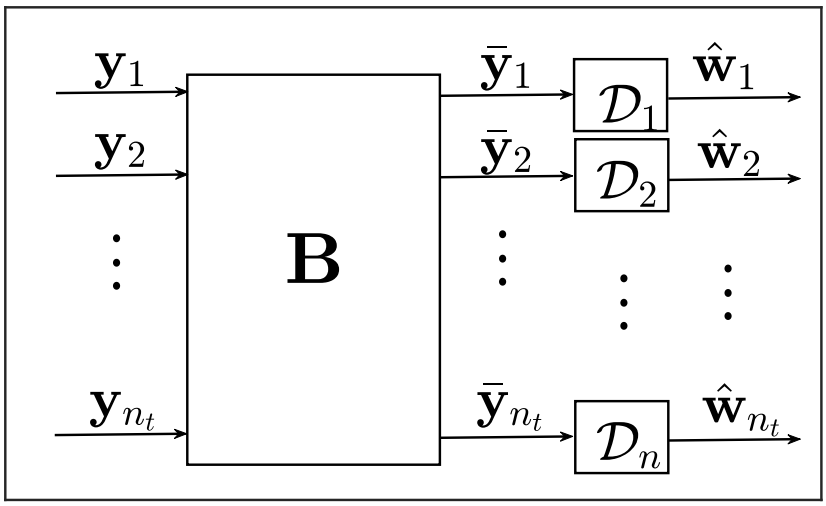

Figure 2.5 – Operation of conventional MIMO linear receivers.

Table 2.2 – Equalization matrix for the ZF and MMSE linear receivers

	Equalization matrix (**B**)
ZF	$\mathbf{H}^{\dagger}$
MMSE	$s\mathbf{H}'\left(\mathbf{I}_{n_t}+s\mathbf{H}\mathbf{H}'\right)^{-1}$

receivers from MMSE linear receivers is the choice of the equalizing matrix **B**. Table 2.2 presents these equalization matrices. The achievable rate for the m-th data stream of these conventional receivers is expressed as,

$$R_{conventional,m}(\mathbf{H},\mathbf{b}_m)=\frac{1}{2}\log\left(\frac{s\left(\mathbf{b}_m'\mathbf{h}_m\right)^2}{\|\mathbf{b}_m\|^2+s\sum_{i\neq m}\left(\mathbf{b}_m'\mathbf{h}_i\right)^2}\right) \tag{2.17}$$

It is possible to observe from (2.16), that upon using these linear receivers the effective noise becomes **BZ**. As matrix **B** elements can have any value, the noise on each data stream may be unevenly amplified, resulting in reduced system efficiency. As an example, lets consider the following 2×2 system,

$$\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix}=\begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{pmatrix}+\begin{pmatrix}\mathbf{z}_1\\ \mathbf{z}_2\end{pmatrix} \tag{2.18}$$

Applying the ZF equalization matrix to the received signals,

$$\begin{pmatrix}1 & -1\\ -1 & 2\end{pmatrix}\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix}=\begin{pmatrix}\mathbf{x}_1\\ \mathbf{x}_2\end{pmatrix}+\begin{pmatrix}\mathbf{z}_1-\mathbf{z}_2\\ -\mathbf{z}_1+2\mathbf{z}_2\end{pmatrix}. \tag{2.19}$$

The effective noise variance that affects the first and second data streams are 2 and 5 respectively.

2.2.2 Integer-Forcing Linear Receivers

The effects of uneven noise amplification are particularly harmful when channel matrices are close to singular. IF linear receivers significantly outperform conventional linear receivers in this scenario, while still having small complexity.

In the IF framework, each transmitter encodes its messages using the same linear lattice codes. Due to this structure, any integer combination of codewords is also a codeword, allowing the receiver to decode these integer combinations with SISO decoders. After decoding a suitable number integer combinations of codewords, it solves a linear system to obtain each individual message. Figure 2.6 illustrates the operation described. Each mes-

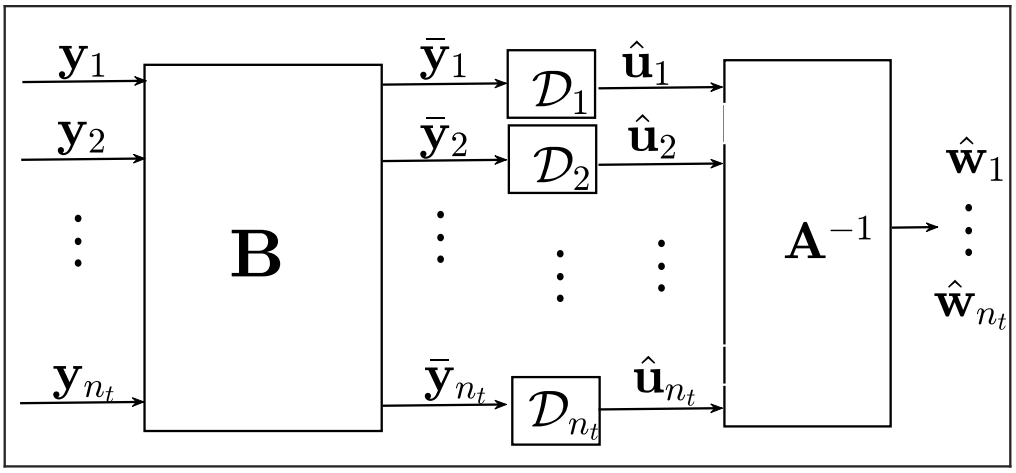

Figure 2.6 – Operation of IF linear receivers.

sage is encoded by a nested lattice code $\mathscr{C}$, producing n_t row vectors in $\mathscr{C} \subset \mathbb{R}^{1\times l}$. Random dithers $\mathbf{d}_m \in \mathbb{R}^{1\times l}$ and the encoded messages $\mathbf{t}_m$ are used to calculate the signal vectors $\mathbf{x}_m$ as,

$$\mathbf{x}_m = [\mathbf{t}_m - \mathbf{d}_m] \bmod \Lambda_c. \tag{2.20}$$

Dither $\mathbf{d}_m$ is a uniformly distributed and independent of $\mathbf{t}_m$, and is known by the transmitters and receiver. It is used to remove dependency between the signal vector and the effective noise. The operator $[\mathbf{r}] \bmod \Lambda_c$ represents the quantization error of $\mathbf{r}$ with respect to the lattice Λ_C,

$$[\mathbf{r}] \bmod \Lambda_c = \mathbf{r} - \arg\min_{\mathbf{t}\in\Lambda_c} \|\mathbf{r}-\mathbf{t}\|. \tag{2.21}$$

Due to the crypto lemma [34], the signals $\mathbf{x}_m$ are statistically independent of $\mathbf{t}_m$.

Upon receiving $\mathbf{Y}$ an equalization matrix $\mathbf{B}$ is applied. This matrix attempts to project the equalized channel $\mathbf{BH}$ onto a non-singular integer (or Gaussian integer if complex modulation is taken into account) matrix $\mathbf{A}$ with $n_t \times n_t$ dimensions. Matrix $\mathbf{A}$ represents the linear combinations of codewords that will be decoded. Furthermore, the receiver also

removes combinations of dithers from the desired signal,

$$\bar{\mathbf{Y}} = [\mathbf{BY} + \mathbf{AD}] \bmod \Lambda_c \tag{2.22}$$

$$= [\mathbf{BHX} + \mathbf{BZ} + \mathbf{AD}] \bmod \Lambda_c \tag{2.23}$$

$$= [\mathbf{AX} + (\mathbf{BH} - \mathbf{A})\mathbf{X} + \mathbf{BZ} + \mathbf{AD}] \bmod \Lambda_c \tag{2.24}$$

$$= [\mathbf{AT} + (\mathbf{BH} - \mathbf{A})\mathbf{X} + \mathbf{BZ}] \bmod \Lambda_c \tag{2.25}$$

Here $\mathbf{D}$ is a matrix in which each row represents the dither sequences $\mathbf{d}_m$. Each row of T represents the encoded messages $\mathbf{t}_m$. The effective noise is $[(\mathbf{BH} - \mathbf{A})\mathbf{X} + \mathbf{BZ}] \bmod \Lambda_c$, and the desired signal component is calculated as $[\mathbf{AT}] \bmod \Lambda_c$. Due to the use of dither, it is possible to observe that these components of the received signal are independent. Is is also important to note that the effective noise is the summation of two undesired effects, noise amplification and the mismatch between the equalized channel $\mathbf{BH}$ and its integer approximation $\mathbf{A}$. Matrices $\mathbf{A}$ and $\mathbf{B}$ have to be carefully chosen in order to minimize both effects. In particular, along the m-th data stream the effective noise variance is expressed as,

$$g_m(\mathbf{a}_m, \mathbf{b}_m) = s\|\mathbf{H}'\mathbf{b}_m - \mathbf{a}_m\|^2 + \|\mathbf{b}_m\|^2, \tag{2.26}$$

where $\mathbf{a}_m$ and $\mathbf{b}_m$ denote the m-th rows of matrices $\mathbf{A}$ and $\mathbf{B}$, respectively. The transmission rate of any single data streams, denoted computational rate, can be calculated as [5] [1],

$$R_{comp,m}(\mathbf{H}, \mathbf{A}, \mathbf{B}) = \log^+\left(\frac{s}{s\|\mathbf{H}'\mathbf{b}_m - \mathbf{a}_m\|^2 + \|\mathbf{b}_m\|^2}\right). \tag{2.27}$$

The achievable sum-rate of IF linear receivers is obtained as,

$$R_{IF} = n_t \max_{\substack{\mathbf{A} \\ Rank(\mathbf{A})=n_t}} \max_{\mathbf{B}} \min_{m=1,\ldots,n_t} R_{comp,m}(\mathbf{H}, \mathbf{A}, \mathbf{B}). \tag{2.28}$$

The optimal equalization matrix $\mathbf{B}_{opt}$, that solves a part of problem (2.28), can be derived as a function of the integer matrix $\mathbf{A}$ and is expressed as [5]

$$\mathbf{B}_{opt} = s\mathbf{A}\mathbf{H}'\left(\mathbf{I}_{n_t} + s\mathbf{H}\mathbf{H}'\right)^{-1}. \tag{2.29}$$

Given the eigendecomposition and Cholesky decomposition,

$$\mathbf{V}\mathbf{D}\mathbf{V}' = \mathbf{I}_{n_t} + s\mathbf{H}'\mathbf{H}, \tag{2.30}$$

$$\mathbf{L}'\mathbf{L} = \left(\mathbf{I}_{n_t} + s\mathbf{H}'\mathbf{H}\right)^{-1}, \tag{2.31}$$

[1] The transmission rate equations presented are suited for complex channels. For real channels, these equations should be multiplied by a factor of $1/2$.

We define matrix $\mathbf{Q}$ in two different ways,

$$\mathbf{Q} = \begin{cases} \mathbf{L}, & \text{or} \\ \mathbf{D}^{-1/2}\mathbf{V}'. \end{cases} \tag{2.32}$$

We use each definition of $\mathbf{Q}$ in different subsection of this work. The freedom to choose between these representations greatly simplifies our analysis. The achievable sum rate can be simplified to,

$$R_{IF} = n_t \max_{\substack{\mathbf{A} \\ Rank(\mathbf{A})=n_t}} \min_{m=1,\dots,n_t} R_{comp,m}(\mathbf{H},\mathbf{A}), \tag{2.33}$$

in which $R_{comp,m}(\mathbf{H},\mathbf{a}_m)$ is expressed as,

$$R_{comp,m}(\mathbf{H},\mathbf{A}) = \log^+\left(\|\mathbf{Q}\mathbf{a}'_m\|^{-2}\right). \tag{2.34}$$

The optimization problem from (2.33) can be written as,

$$\mathbf{A}_{opt} = arg \min_{\substack{\mathbf{A} \\ Rank(\mathbf{A})=n_t}} \max_{m=1,\dots,n_t} \|\mathbf{Q}\mathbf{a}'_m\|^2, \tag{2.35}$$

The search for the optimal $\mathbf{A}$ in (2.35) is equivalent to the search for the shortest set of n linearly independent vectors (SIVP) in the lattice generated by $\mathbf{Q}$ in any of its two forms. Although SIVP is considered to be an NP-hard problem, lattice reduction algorithms such as LLL and BKZ can be used to obtain an approximate answer in polynomial time with a small penalty to capacity. In [18] and [17], an in-depth analysis on the performance of IF receivers with the use of lattice reduction techniques was performed by the authors.

2.2.2.1 Compute and Forward Transform (CFT)

In [35] the authors present a variation to the IF framework in which the receiver uses successive interference cancellation (SIC) to increase the achievable sum rate. By decoding the combination of codewords in order of decreasing computational rate and using SIC, the receiver can remove the combinations with higher computational rates from the combinations with lower computational rates. By doing so, the different streams of data are no longer limited by the worst computational rate as in the traditional IF framework, which allows each transmitter to encode its messages with different rates. The sum rate of the C-FT is,

$$R_{sum_{C-FT}} = \sum_{m=1}^{n_t} R_{comp,opt,m}(\mathbf{H},\mathbf{A}). \tag{2.36}$$

Here $R_{comp,opt,m}(\mathbf{H},\mathbf{A})$ represents the maximum computational rate with respect to $\mathbf{A}$. Overall C-FT is more complex than IF due to the use of SIC. However its attainable sum rate is equal or greater than IF receiver's sum rate $\left(R_{sum_{C-FT}} \leq R_{IF}\right)$ as,

$$n_t \min_{m=1,\dots,n_t} R_{comp,opt,m}(\mathbf{H},\mathbf{A}) \leq \sum_{m=1}^{n_t} R_{comp,opt,m}(\mathbf{H},\mathbf{A}). \tag{2.37}$$

2.2.2.2 Successive IF

More recently in [16] by using receivers that perform noise prediction, the authors introduced the Successive Integer-Forcing (SIF) which was proven to attain MIMO capacity for uniform power allocation systems,

$$R_{sum_{SIF}} = \log\left(\det\left(\mathbf{I}_{n_t} + s\mathbf{H}'\mathbf{H}\right)\right). \tag{2.38}$$

It is important to note that, even though CFT and SIF outperform the traditional IF linear receiver, both systems require that each of its transmitters allocate different transmission rates. This is only possible if either the receiver feedbacks this information to the transmitters or if all transmitters have complete channel state information (CSI).

Chapter 3

Approximate Sum Rate for Integer-Forcing Receiver

3.1 Introduction

In this chapter we start presenting our contributions. The contents of it were published in [36]. For simplicity, only systems with an equal number of transmitting and receiving antennas (n) are considered in this chapter. We derived two approximations to the achievable sum-rate of the IF linear receiver and their respective analytical probability density functions (PDF). The first approximation is based on the Gauss-Lagrange algorithm for real or complex systems with two antennas at the transmitter and receiver (2×2 arrays) and uncorrelated channels. The second approximation considers an $n \times n$ array for both correlated and uncorrelated channels and its derivation is based on *Minkowiski*'s second theorem. In order to validate our approximations, three metrics were evaluated, the sum-rate's mean value, the outage probability, and the outage rate.

It is interesting to note that the IF subject inherits many concepts from MIMO theory subject to Rayleigh fading, which has been fully analyzed using random matrix theory. Unfortunately, there is a scarcity of results joining the areas of random matrix and lattice theory. Although there are several famous random ensembles of lattices, such as the Minkowski-Hlawka-Siegel (MHS) ensemble, the Rogers ensemble, and random construction A (Loeliger's ensemble) [37], none of them are useful for this analysis. Since IF linear receivers rely on lattices whose generator matrix is a function of independent and identical distributed channel gains, these lattices do not satisfy the criteria (such as fixed cell volume) to belong to any of these ensembles. As a consequence, there is no closed-form expression to obtain the ergodic achievable sum-rate of the IF linear receiver and therefore, it is necessary to compute it through a series of Monte-Carlo simulations, requiring an NP-complete optimization procedure at every iteration. Our approximations are computable more easily since the PDF for the approximation to the $n \times n$ array are given as one fold integral and its mean value can be computed as a simple function of the derivative of the characteristic function.

The remainder of the chapter is organized as follows. We begin Section 2.2 by presenting our approximations. In section 2.3, the PDFs of our approximations are derived. In Section 2.4, we compare simulation results to those obtained from numerical calculation of our approximations. Finally, Section 2.5 concludes the chapter.

3.2 Proposed Approximations

In this section, we present the largest Minkowski Minima, λ_n, for the IF linear receiver and then the approximation for the 2×2 and $n \times n$ arrays.

3.2.1 IF and the Minkowski Minima

It is possible to observe, from equations (2.33) and (2.35), that the calculation of the IF achievable sum rate involves a complex optimization problem. Since there is no closed form expression to determine the optimal matrix $\mathbf{A}$, it is difficult to obtain the distribution of R_{IF} in (2.33). Given $\mathbf{Q}$, as the basis for a lattice $\Lambda(\mathbf{Q})$, the calculation of the optimal matrix $\mathbf{A}$ in problem (2.33) can also be viewed as the search for the unimodular matrix that will result in a new lattice basis $\bar{\mathbf{Q}} = \mathbf{QA}$ with smallest possible vector norms. The computable IF rate (2.34) can then be expressed as a function of $\bar{\mathbf{Q}}$,

$$R_{comp,m} = \log^+\left(\|\mathbf{Q}\mathbf{a}'_m\|^2\right) \tag{3.1}$$

$$= \log^+\left(\|\bar{\mathbf{q}}_m\|^{-2}\right), \tag{3.2}$$

in which $\bar{\mathbf{q}}_m$ is the m-th column of $\bar{\mathbf{Q}}$. The vector $\bar{\mathbf{q}}_m$ that minimizes (3.2) is the one with the largest norm. Since $\bar{\mathbf{Q}}$ is the representation of $\Lambda(\mathbf{Q})$ with the smallest possible basis norms, the vector $\bar{\mathbf{q}}_m$ with largest norm is equivalent to the n-th successive lattice minima ($\lambda_n(\mathbf{Q})$) of $\Lambda(\mathbf{Q})$. Therefore the IF achievable sum rate can also be expressed as,

$$R_{IF} = n \log^+\left(\lambda_n(\mathbf{Q})^{-2}\right). \tag{3.3}$$

In the next subsections, we derive approximations to $\lambda_n(\mathbf{Q})$. For each approximation, we are able to obtain their distribution. These are used to derive the approximations to the IF sum rate.

3.2.2 The Approximation for 2×2 array

For a real two-dimensional lattice, its reduction problem (2.35) can be solved by the Gauss-Lagrange algorithm. However, obtaining the distribution of the output of the

algorithm is excessively complex. Fortunately, if the lattice basis $\mathbf{Q} = \mathbf{L}$ is used for the channels considered in this section, that is, real or complex uncorrelated Gaussian distributed, it has been observed that 98% of the channel realizations requires only one iteration of the Gauss-Lagrange algorithm to obtain $\lambda_2(\mathbf{L})$, as will be shown in section 3.4. Therefore, our approximation is based on the output of one iteration of algorithm 1.

3.2.2.1 Analysis of the Gauss-Lagrange Algorithm

Given matrix $\mathbf{L}$ defined in (2.31) as input of the algorithm, consider $\mathbf{l}_m$ as the m-th column and l_{ab} as the element at position (a, b). The result ($\bar{\mathbf{L}}_1$) of one iteration of the Gauss-Lagrange lattice reduction is,

$$\bar{\mathbf{L}}_1 = \begin{cases} (\mathbf{l}_1, \mathbf{t}_1), & \text{if } ||\mathbf{l}_1|| < ||\mathbf{l}_2|| \& ||\mathbf{l}_1|| < ||\mathbf{t}_1||, \\ (\mathbf{t}_1, \mathbf{l}_1 - \lceil \mu(\mathbf{t}_1, \mathbf{l}_1) \rfloor \mathbf{t}_1), & \text{if } ||\mathbf{l}_1|| < ||\mathbf{l}_2|| \& ||\mathbf{l}_1|| > ||\mathbf{t}_1||, \\ (\mathbf{l}_2, \mathbf{t}_2), & \text{if } ||\mathbf{l}_2|| < ||\mathbf{l}_1|| \& ||\mathbf{l}_2|| < ||\mathbf{t}_2||, \\ (\mathbf{t}_2, \mathbf{l}_2 - \lceil \mu(\mathbf{t}_2, \mathbf{l}_2) \rfloor \mathbf{t}_2), & \text{if } ||\mathbf{l}_2|| < ||\mathbf{l}_1|| \& ||\mathbf{l}_2|| > ||\mathbf{t}_2||. \end{cases} \tag{3.4}$$

In which,

$$\mu(\mathbf{a}, \mathbf{t}) = \frac{\mathbf{a}' \cdot \mathbf{t}}{||\mathbf{a}||^2}, \tag{3.5}$$

$$\mathbf{t}_1 = \mathbf{l}_2 - \lceil \mu(\mathbf{l}_1, \mathbf{l}_2) \rfloor \mathbf{l}_1, \tag{3.6}$$

$$\mathbf{t}_2 = \mathbf{l}_1 - \lceil \mu(\mathbf{l}_2, \mathbf{l}_1) \rfloor \mathbf{l}_2. \tag{3.7}$$

Since all possible outcomes of $\bar{\mathbf{L}}_1$ are also representations of $\Lambda(\mathbf{L})$, we use the minimum value among the largest column norm of each possible outcome of $\bar{\mathbf{L}}_1$ in order to approximate $\lambda_2(\mathbf{L})$. To further simplify the analysis, the rounding operation is removed, so that,

$$\begin{aligned} \lceil \mu(\mathbf{l}_1, \mathbf{l}_2) \rfloor &\approx \mu(\mathbf{l}_1, \mathbf{l}_2) = l_{12}/l_{11}, \\ \lceil \mu(\mathbf{l}_2, \mathbf{l}_1) \rfloor &\approx \mu(\mathbf{l}_2, \mathbf{l}_1) = \frac{l_{11} l_{12}^*}{|l_{12}|^2 + l_{22}^2}. \end{aligned} \tag{3.8}$$

Notice that since $\mathbf{L}$ is the result of a Cholesky decomposition, $l_{21} = 0$, l_{11} and l_{22} are real and positive, l_{12} may be real or complex following the format of the elements of $\mathbf{H}$. Using (3.8), $\mathbf{t}_1$ and $\mathbf{t}_2$ are approximated as,

$$\begin{aligned} \mathbf{t}_1 &\approx \bar{\mathbf{t}}_1 = \mathbf{l}_2 - \mu(\mathbf{l}_1, \mathbf{l}_2)\mathbf{l}_1 = (0, l_{22})^T \\ \mathbf{t}_2 &\approx \bar{\mathbf{t}}_2 = \mathbf{l}_1 - \mu(\mathbf{l}_2, \mathbf{l}_1)\mathbf{l}_2 = \left(\frac{l_{11} l_{22}^2}{l_{22}^2 + |l_{12}|^2}, -\frac{l_{11} l_{22} l_{12}^*}{l_{22}^2 + |l_{12},|^2} \right)^T. \end{aligned} \tag{3.9}$$

Removing the rounding operations and using (3.9), we have,

$$\begin{aligned} \lceil \mu(\mathbf{t}_1, \mathbf{l}_1) \rfloor &\approx \mu(\bar{\mathbf{t}}_1, \mathbf{l}_1) = 0, \\ \lceil \mu(\mathbf{t}_2, \mathbf{l}_2) \rfloor &\approx \mu(\bar{\mathbf{t}}_2, \mathbf{l}_2) = 0. \end{aligned} \tag{3.10}$$

Considering the possible outcomes of $\bar{\mathbf{L}}_1$ and their respective restrictions, along with the equations (3.8), (3.9), (3.10), the proposed approximation to $\lambda_2(\mathbf{L})$ is expressed as,

$$\lambda_2(\mathbf{L}) \approx \hat{\lambda}_2(\mathbf{L}) = \min(\tau_1, \tau_2), \tag{3.11}$$

in which

$$\begin{aligned} \tau_1 &= \max(l_{11}, l_{22}), \\ \tau_2 &= \max\left(||\mathbf{l}_2||, \frac{l_{11} l_{22}}{\sqrt{|l_{12}|^2 + l_{22}^2}} \right). \end{aligned} \tag{3.12}$$

The approximation to the 2×2 IF sum rate is defined as,

$$\hat{R}_{IF_2} = 4 \log^+ \left(\hat{\lambda}_2^{-1} \left(\mathbf{L}' \right) \right). \tag{3.13}$$

3.2.3 The Approximation for $n \times n$ Array

Algorithms that perform lattice reduction on dimensions greater then two are significantly more complex than the Gauss-Lagrange algorithm, and therefore, we are unable to replicate the same reasoning to obtain a good approximation to the IF sum rate for arrays with an arbitrary number of antennas. Using a different approach to the problem, we present, in this subsection, an approximation to the $n \times n$ array IF sum rate that is tight at high SNR.

We begin proposing an approximation to $\lambda_n(\mathbf{Q})$ expressed as,

$$\lambda_n(\mathbf{Q}) \approx \sqrt{\gamma_n} \det(\mathbf{Q}'\mathbf{Q})^{1/2n}. \tag{3.14}$$

It is important to note that the values of the Hermite constant γ_q in table 2.1 are for real valued lattices. For complex valued lattices, $n = 2q$ should be considered.

The rationale for the approximation comes from the upper bound to the geometrical mean of Minkowski minima, given by,

$$\sqrt{\gamma_n} \det(\mathbf{Q}'\mathbf{Q})^{1/2n} \geq \left(\prod_{i=1}^{n} \lambda_i(\mathbf{Q}) \right)^{\frac{1}{n}}. \tag{3.15}$$

Since the geometrical mean of a set is always smaller than its largest element,

$$\lambda_n(\mathbf{Q}) \geq \left(\prod_{i=1}^{n} \lambda_i(\mathbf{Q}) \right)^{\frac{1}{n}}, \tag{3.16}$$

using (3.15) and (3.16), we will suppose that (3.14) is valid. Unfortunately, given the order of the inequalities, we do not have a bound, and that is the reason to propose an approximation. The tighnesss of this approximation will be shown by the analysis of the gap and also by simulations. Using (3.14), our approximation $\hat{R}_{IF_n}$, to the achievable sum rate R_{IF} will be given by,

$$\hat{R}_{IF_n} = n \log^+ \left(\left(\gamma_n \det(\mathbf{Q}'\mathbf{Q})^{1/n} \right)^{-1} \right). \tag{3.17}$$

3.2.3.1 Lower Bound to Compute and Forward Transform

Given (2.36), it is possible to express the achievable C-FT sum rate as,

$$R_{sum_{C-FT}} = \log^{+}\left(\prod_{m=1}^{n}\lambda_m^{-2}(\mathbf{Q})\right). \tag{3.18}$$

Due to (3.15), it is possible to observe that $\hat{R}_{IF_n} \leq R_{sum_{C-FT}}$, and therefore our approximation to the $n \times n$ array IF sum rate also serves as a lower bound to the C-FT sum rate.

3.2.4 The Error of the $n \times n$ Array Approximation

In order to assess the $n \times n$ array approximation error as the SNR grows, let us compute the coefficient of variation defined as,

$$cv(s) = \frac{\sigma_e}{\mu_e} = \frac{\sqrt{\mathbb{E}\left[\left(R_{IF}(s) - \hat{R}_{IF}(s)\right)^2\right]}}{\mathbb{E}\left[R_{IF}(s)\right]}. \tag{3.19}$$

In which σ_e represents the standard deviation of the error. The coefficient of variation shows the extent of variability of the error in relation to the mean sum rate, and therefore is a metric that allows the comparison of the approximation error between different arrays. Using the definition of $\mathbf{Q} = \mathbf{D}(s)^{-1/2}\mathbf{V}'$, σ_e^2 can be expressed as,

$$\sigma_e^2 = \mathbb{E}\left[\left(n\log\left(\frac{\gamma_n \det(\mathbf{D}(s)^{-1})^{1/n}}{\lambda_n^2(\mathbf{D}(s)^{-1/2}\mathbf{V}')}\right)\right)^2\right]. \tag{3.20}$$

At high SNR, (2.30) can be written as $\mathbf{V}\mathbf{D}\mathbf{V}' \approx s\mathbf{H}'\mathbf{H}$, and therefore,

$$\sigma_e^2 \approx \mathbb{E}\left[\left(n\log\left(\frac{\gamma_n \det(s^{-1}\boldsymbol{\Delta}^{-1})^{1/n}}{\lambda_n^2\left(s^{-1/2}\boldsymbol{\Delta}^{-1/2}\mathbf{V}'\right)}\right)\right)^2\right], \tag{3.21}$$

where the matrix $\boldsymbol{\Delta}$ is the eigenvalue matrix of $\mathbf{H}'\mathbf{H}$. Now using the property that $\lambda_n(c\mathbf{L}) = c\lambda_n(\mathbf{L})$ and $\det(c\mathbf{L}) = c^n\det(\mathbf{L})$ we finally have,

$$\sigma_e^2 \approx \mathbb{E}\left[\left(n\log\left(\frac{\gamma_n \det(\boldsymbol{\Delta}^{-1})^{1/n}}{\lambda_n^2\left(\boldsymbol{\Delta}^{-1/2}\mathbf{V}'\right)}\right)\right)^2\right]. \tag{3.22}$$

When SNR is large, it is possible to observe that error variance σ_e^2 is constant. Since μ_e increases as the SNR grows, it is possible to conclude that,

$$\lim_{s\to\infty} cv(s) = 0. \tag{3.23}$$

In section 3.4, we will show through a series of numerical simulations how cv behaves for a limited SNR as the number of antennas increases.

3.2.5 Approximate Degrees of Freedom of the $n \times n$ Array Approximation

The approximate degrees of freedom (η) of a certain coding scheme is defined as [38],

$$\eta = \lim_{s\to\infty} \frac{\hat{R}_{\text{sum}}}{\log(s)}, \tag{3.24}$$

in which $\hat{R}_{sum}(s)$ represents its approximate sum transmission rate. In this subsection we prove that our $n \times n$ array approximation attains n degrees of freedom, the same number as the IF linear receiver (proved in [5]).

As shown in subsection 3.2.4, the approximation can be expressed as,

$$\begin{aligned}\hat{R}_{IF_n} &= n\log^+\left(\left(\gamma_n \det(\mathbf{Q}'\mathbf{Q})^{1/n}\right)^{-1}\right)\\ &= -n\log\left(\gamma_n \det(s^{-1}\boldsymbol{\Delta}^{-1})^{1/n}\right).\end{aligned} \tag{3.25}$$

Therefore, the approximate degrees of freedom for $n \times n$ array approximation is

$$\begin{aligned}\eta &= \lim_{s\to\infty} \frac{-n\log\left(\gamma_n \det(s^{-1}\boldsymbol{\Delta}^{-1})^{1/n}\right)}{\log(s)}\\ &= \lim_{s\to\infty} \frac{-\log\left(\gamma_n \det(\boldsymbol{\Delta})^{-1}\right) - \log(s^{-n})}{\log(s)} = n.\end{aligned}$$

3.3 Distribution of the Approximations

In order to derive the ergodic achievable sum rate, outage probability, and outage sum rate, we will investigate the distribution of the approximations of the 2×2 and $n\times n$ arrays in the following subsections. For the 2×2 array approximation, we will analyze scenarios with real and complex uncorrelated channel gains. For the $n\times n$ array approximation, complex correlated and uncorrelated channel gains will be considered.

3.3.1 Distribution of the 2×2 Array Approximation

First, we derive the joint PDF of the elements of matrix $\mathbf{L}$ and with this result, we obtain the distribution of the 2×2 array approximation. We will work with $n \times n$ arrays and eventually set $n = 2$. We will use $\beta = 1$ to represent the real case, and $\beta = 2$ to represent the complex case.

$$F_{\hat{\lambda}_{2r}}(u) = \int_0^1 \int_{-1}^1 \int_0^1 p_{jointr}(l_{11}, l_{12}, l_{22}) \times \left(H(u-l_{11})H(u-l_{22}) + H\left(u-\sqrt{l_{12}^2+l_{22}^2}\right)H\left(u-\frac{l_{11}l_{22}}{\sqrt{l_{12}^2+l_{22}^2}}\right) - H(u-l_{11})H\left(u-\sqrt{l_{12}^2+l_{22}^2}\right)H\left(u-\frac{l_{11}l_{22}}{\sqrt{l_{12}^2+l_{22}^2}}\right)H(u-l_{22})\right) \mathrm{d}l_{22}\mathrm{d}l_{12}\mathrm{d}l_{11}. \quad (3.27)$$

Theorem 3.1. *The joint PDF of the elements of matrix* $\mathbf{L}$ *or* $\mathbf{L}'$ *defined in* (2.31) *for an* $n \times n$ *array is,*

$$\mathscr{P}_{\mathbf{L}}(\mathbf{L}) = \frac{2^n C_{\beta,n}}{s^{\frac{\beta n^2}{2}}} \exp\left(-\frac{\beta}{2s} tr((\mathbf{L}'\mathbf{L})^{-1} - \mathbf{I}_n)\right) \times \det((\mathbf{L}'\mathbf{L})^{-1} - \mathbf{I}_n)^{\frac{\beta}{2}-1} \det(\mathbf{L}'\mathbf{L})^{-\beta(n-1)-2} \times \prod_{j=1}^{n} l_{jj}^{\beta(n-j)+1} H((\mathbf{L}'\mathbf{L})^{-1} - \mathbf{I}_n). \quad (3.26)$$

The proof of Theorem 3.1 is presented in Appendix A.1.

Setting $n = 2$, we now separate the analysis of the real and complex cases in order to obtain the cumulative distribution function (CDF) of $\hat{\lambda}_2$ for both scenarios.

3.3.1.1 Real Case

Theorem 3.2. *The joint CDF of the approximation of the second Minkowski minima for the* 2×2 *real array* $\hat{\lambda}_{2r}$ *is expressed in* (3.27).

The proof of Theorem 3.2 is presented in Appendix A.2.

3.3.1.2 Complex Case

Theorem 3.3. *The joint CDF of the approximation of the second Minkowski minima for the* 2×2 *complex array* $\hat{\lambda}_{2c}$ *is presented in* (3.28). *Here the constant R is defined as,*

$$R = \sqrt{1 - l_{11}^2 - l_{22}^2 + l_{11}^2 l_{22}^2}. \quad (3.29)$$

The proof of Theorem 3.3 is presented in Appendix A.3.

3.3.1.3 PDF of $\hat{R}_{IF_2}$

The PDF $p_{\hat{\lambda}_2}(u)$ for both cases is obtained as the derivative of the appropriate CDF ($F_{\hat{\lambda}_{2r}}(u)$ for the real case or $F_{\hat{\lambda}_{2c}}(u)$ for the complex case). Finally the PDF of $\hat{R}_{IF_2}$ given

$$F_{\hat{\lambda}_{2c}}(u) = \int_0^1 \int_{-1}^1 \int_0^1 \int_{-R}^R p_{jointc}(l_{11}, l_{12}, l_{22}, k_{12}) \times \left(H\left(u - \sqrt{l_{12}^2 + l_{22}^2 + k_{12}^2}\right) H\left(u - \frac{l_{11} l_{22}}{\sqrt{l_{12}^2 + l_{22}^2 + k_{12}^2}}\right) + H(u - l_{22}) H(u - l_{11}) - H\left(u - \sqrt{l_{12}^2 + l_{22}^2 + k_{12}^2}\right) H\left(u - \frac{l_{11} l_{22}}{\sqrt{l_{12}^2 + l_{22}^2 + k_{12}^2}}\right) H(u - l_{11}) H(u - l_{22}) \right) \mathrm{d}k_{12} \mathrm{d}l_{22} \mathrm{d}l_{12} \mathrm{d}l_{11}. \tag{3.28}$$

in (3.13) can be obtained as,

$$p_{\hat{R}_{IF_2}}(r) = \begin{cases} 0, & \text{if } r < 0, \\ K_2 \delta_{dirac}(r) - \frac{2^{-r/4}}{4\log(e)} p_{\hat{\lambda}_2}\left(2^{-r/4}\right), & \text{otherwise,} \end{cases} \tag{3.30}$$

in which $\delta_{dirac}(.)$ is the dirac delta function and the constant K_2 is given by,

$$K_2 = \int_{-\infty}^0 -\frac{2^{-r/4}}{4\log(e)} p_{\hat{\lambda}_2}\left(2^{-r/4}\right) dr. \tag{3.31}$$

3.3.2 Distribution of the $n \times n$ Array Approximation

Since

$$\mathbf{Q}'\mathbf{Q} = \mathbf{V}\mathbf{D}^{-1}\mathbf{V}', \tag{3.32}$$

then applying (2.30) in (3.17), the following can be written

$$\hat{R}_{IF_n} = n \log^+\left(\left(\gamma_n \det\left(\mathbf{I}_n + s\mathbf{H}'\mathbf{H}\right)^{-1/n}\right)^{-1}\right). \tag{3.33}$$

This can be expanded as

$$\hat{R}_{IF_n} = \max\left(0, \log(\gamma_n^{-n}) + \log\left(\det\left(\mathbf{I}_n + s\mathbf{H}'\mathbf{H}\right)\right)\right) \tag{3.34}$$

$$= \max\left(0, \log(\gamma_n^{-n}) + C_{\text{MIMO}}\right), \tag{3.35}$$

where we define

$$C_{\text{MIMO}} = \log\left(\det\left(\mathbf{I}_n + s\mathbf{H}'\mathbf{H}\right)\right). \tag{3.36}$$

Note that the approximate achievable sum rate is the classical MIMO capacity with uniform power allocation [39] plus a constant. Moreover, the distribution of (3.36) is given as a function of the characteristic function as [40],

$$p_{\text{MIMO}}(r) = \frac{1}{2\pi} \int_R e^{\imath\omega r} \Phi(\omega) d\omega. \tag{3.37}$$

Here the characteristic function $\Phi(\omega)$, for the uncorrelated channel, is given as [40],

$$\Phi_{un}(\omega) = K \det(hu_{j,k})_{j,k=1,\dots,n}, \tag{3.38}$$

with

$$K^{-1} = \frac{1}{n!}\prod_{j=1}^{n}\Gamma(j+1)\Gamma(j+2n), \tag{3.39}$$

$$hu_{j,k} = s^{-\imath\omega}\Gamma(j+k-1)U(\imath\omega, \imath\omega - j - k + 2; s^{-1}), \tag{3.40}$$

where $U(a,b;c)$ is the Tricomi confluent hypergeometric function, and $\Gamma(.)$ defines the gamma function.

In the same way, $\Phi(\omega)$ is given, for the correlated channel, as

$$\Phi_{cor}(\omega) = K_{\Sigma} \det(hc_{j,k})_{j,k=1,\dots,n}, \tag{3.41}$$

in which,

$$K_{\Sigma} = \frac{-n!}{\Gamma(n+1)\prod_{j>k}(\sigma_j - \sigma_k)}\prod_{j=1}^{n}\frac{\sigma_j^{-n}}{\Gamma(j)} \tag{3.42}$$

$$hc_{j,k} = \sigma_j^{k-\imath\omega}s^{-\imath\omega}\Gamma(k)U\left(\imath\omega, \imath\omega - k + 1; s^{-1}\sigma_j^{-1}\right). \tag{3.43}$$

The variables σ_i are the eigenvalues of the covariance matrix $\mathbf{\Sigma}$. Now, since the approximate achievable sum rate is a linear function of C_{MIMO}, the PDF of the approximation can be calculated as,

$$p_{\hat{R}_{IF_n}}(r) = \begin{cases} 0, & \text{if } r < 0, \\ K_n\delta_{dirac}(r) + p_{\text{MIMO}}\left(r + \log(\gamma_n^n)\right), & \text{otherwise.} \end{cases} \tag{3.44}$$

Here K_n is given by,

$$K_n = \int_{-\infty}^{0} p_{\text{MIMO}}\left(r + \log(\gamma_n^n)\right) dr. \tag{3.45}$$

3.3.2.1 Successive Integer Forcing Distribution

As mentioned in 2.2.2.2, SIF receivers are proven to attain MIMO capacity. Therefore the PDF of its achievable sum rate can be calculated through (3.37), using the characteristic functions (3.38) and (3.41) for uncorrelated and correlated channels respectively.

3.3.3 Ergodic Capacity, Outage and Outage Rate of the Approximations

Considering $\hat{R}_{IF}$ a general notation which represents either $\hat{R}_{IF_2}$ or $\hat{R}_{IF_n}$, the mean value of the approximations can be calculated with the following integral,

$$C_{\hat{R}_{IF}} \triangleq \mathbb{E}\left(\hat{R}_{IF}\right) = \int_{-\infty}^{\infty} r p_{\hat{R}_{IF}}(r) dr. \tag{3.46}$$

However for R_{IF_n}, when the characteristic function of C_{MIMO} exists around 0, it is simpler to obtain the mean value as a function of the derivative of its characteristic function,

$$C_{\hat{R}_{IF_n}} = \log\left(\gamma_n^n\right) - \imath \frac{d\Phi(\omega)}{d\omega}\bigg|_{\omega=0}. \tag{3.47}$$

The ergodic capacity is important for the case in which limited feedback from the receiver is possible and when blocklength is much larger than the coherence period. Under this operation, the system attains sum-rate and the ergodic capacity represents the average transmission rate over different channel realizations. In the absence of feedback the figure of merit is the outage rate, defined as,

$$R_{outage}(\rho) = \sup\left(R : p_{outage}(R) \leq \rho\right), \tag{3.48}$$

in which p_{outage} is the outage probability defined as,

$$p_{outage}(R) = \mathbb{P}\left(\hat{R}_{IF}\left(\mathbf{H}\right) \leq R\right). \tag{3.49}$$

The parameter $\hat{R}_{IF}\left(\mathbf{H}\right)$ represents the approximate achievable rate of the receiver for a given channel realization $\mathbf{H}$. Having the CDF $P_{\hat{R}_{IF}}(R) = \int_{-\infty}^{R} p_{\hat{R}_{IF}}(r)dr$, the outage rate is obtained as the inverse function of the CDF

$$Outrate(\rho) = P_{\hat{R}_{IF}}^{-1}(\rho). \tag{3.50}$$

The outage rate is important for transmitters in an open-loop operation. In this case, each transmitter calculates its transmission rate, ensuring that the outage probability will be no larger than a value previously set.

3.4 Simulations

In this section, we present several simulations in order to show the accuracy of the proposed approximations. For all numerical simulations, we have considered 10^4 repetitions of the system model described in section 2.2. In order to obtain the matrix $\mathbf{A}$, which maximizes (2.34), the BKZ algorithm presented and implemented in the number theory library (NTL) was used with block size n [41].

3.4.1 Gauss-Lagrange and second Minkowski Minima.

In this simulation, we analyze the frequency of channel realizations in which one iteration of the Gauss-Lagrange attains the optimal output. The results are presented in Figure 3.1 considering real and complex uncorrelated channel gains. It is possible to observe that, for over 98% of channel realizations, only one iteration of the Gauss-Lagrange is needed to obtain λ_2.

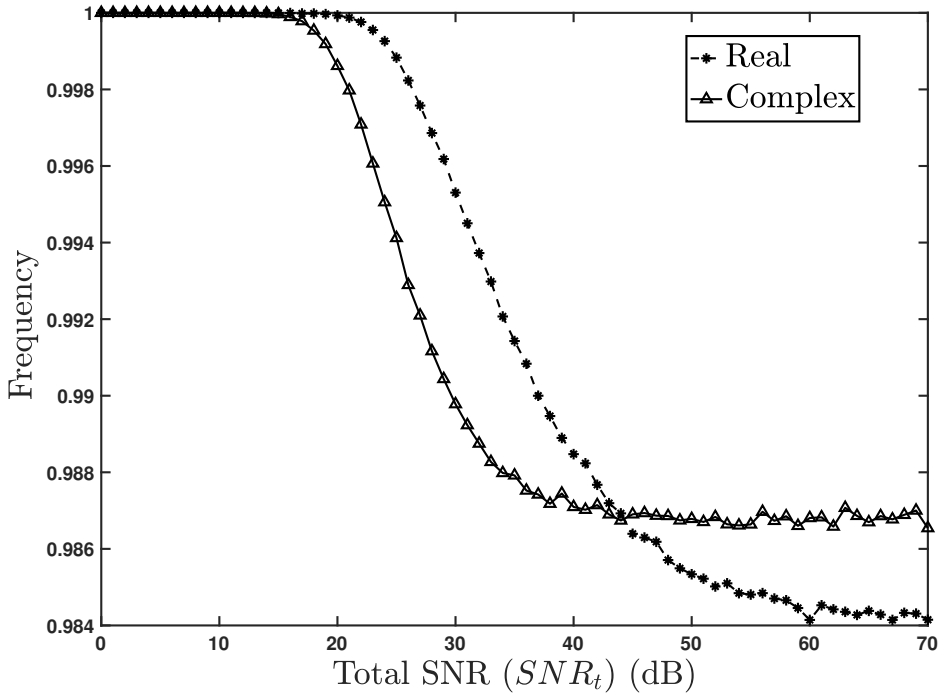

Figure 3.1 – Frequency of channel realizations in which the output of one iteration of Gauss-Lagrange is optimal versus Total SNR (dB)

3.4.2 Mean Root Squared Error per Antenna

In order to investigate the error associated with $\hat{R}_{IF_n}$ as the number of antennas grows, we have performed numerical simulations of the coefficient of variation of the $n \times n$ complex array approximation (3.19) considering $s = 40dB$. Note that, the error was computed using the exact known values of γ_n (as present in Table 2.1) for $n = \{2, 3, 4, 12\}$ and the bound (2.9) for $n = [21, 40]$. Figure 3.2 presents these results. It is possible to observe that for all arrays tested, the coefficient of variation is smaller than 3.5%, and for large arrays it approaches 2.3%.

3.4.3 Ergodic Rate $C_{\hat{R}_{IF}}$

In this subsection, we have compared the ergodic sum-rate defined in (3.46), as a function of the total SNR ($\text{SNR}_t = ns$). Curves representing the ergodic sum-rate for the IF approximation, IF linear receiver, ZF linear receiver, and also the MIMO capacity are presented in all figures.

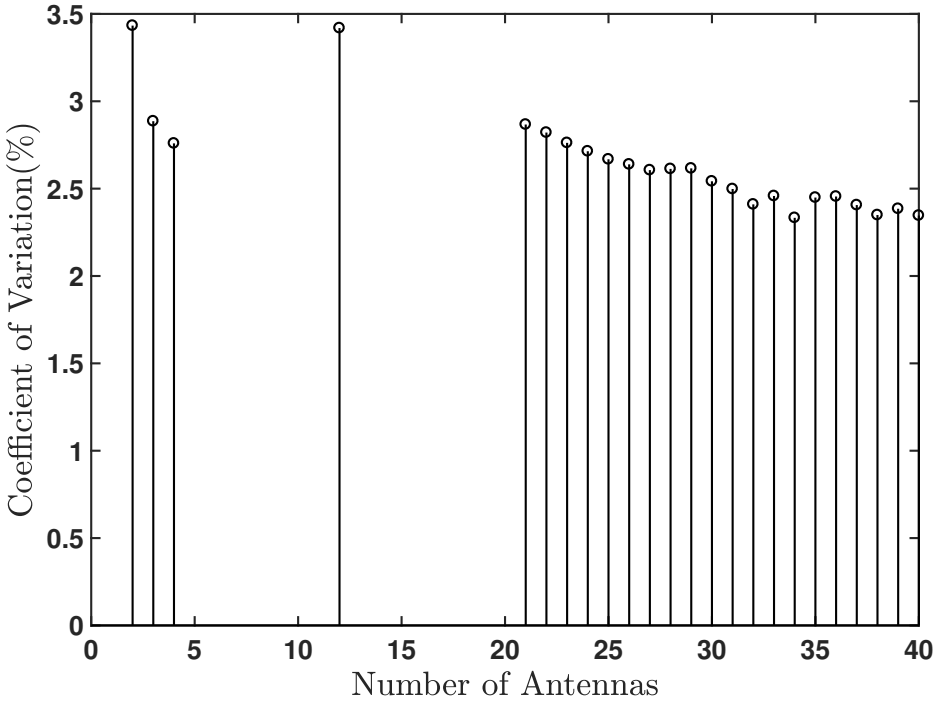

Figure 3.2 – Coefficient of Variation vs Number of antennas

3.4.3.1 Uncorrelated Channels

In Figure 3.3, we analyze the tightness of $\hat{R}_{IF_2}$ for uncorrelated channels with real and complex gains. As can be observed, the simulated R_{IF} and $\hat{R}_{IF_2}$ curves are close for all values of SNR. The performance of $\hat{R}_{IF_n}$ for uncorrelated complex channels is analyzed in Figure 3.4 for setups with 2×2 and 4×4 arrays and in Figure 3.5, for 40×40 array setup (black curves). It is possible to note that the approximation is indeed very tight at high SNR for all the cases.

-

3.4.3.2 Correlated Channels

In Figure 3.5 (blue curves) and 3.6, the simulations consider the existence of correlation among transmitting antennas with covariance matrices $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_4$ and $\mathbf{\Sigma}_{40}$ for the 2×2, 4×4 and 40×40 arrays, respectively. All covariance matrices follow the same construction method, in which all elements are equal to 0.8 with the exception of the elements in the diagonal, which are unit. These matrices were chosen because they represent channels that are highly correlated, contrasting with the previous simulations done for uncorrelated channels. For high SNR values, the curves R_{IF} and $\hat{R}_{IF_n}$ are very close regardless of the number of antennas or correlation among transmitting antennas. We find that, as correlation among antennas grows, the ergodic rate degrades, as expected. However, it is interesting to

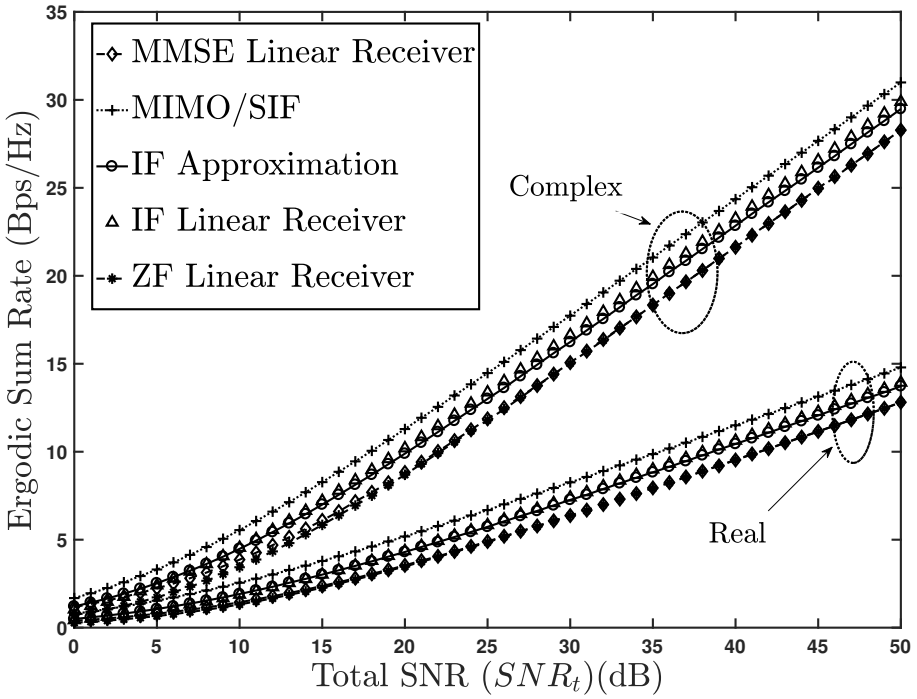

Figure 3.3 – Ergodic Sum Rate versus Total SNR (dB) for 2 × 2 array with uncorrelated channels.

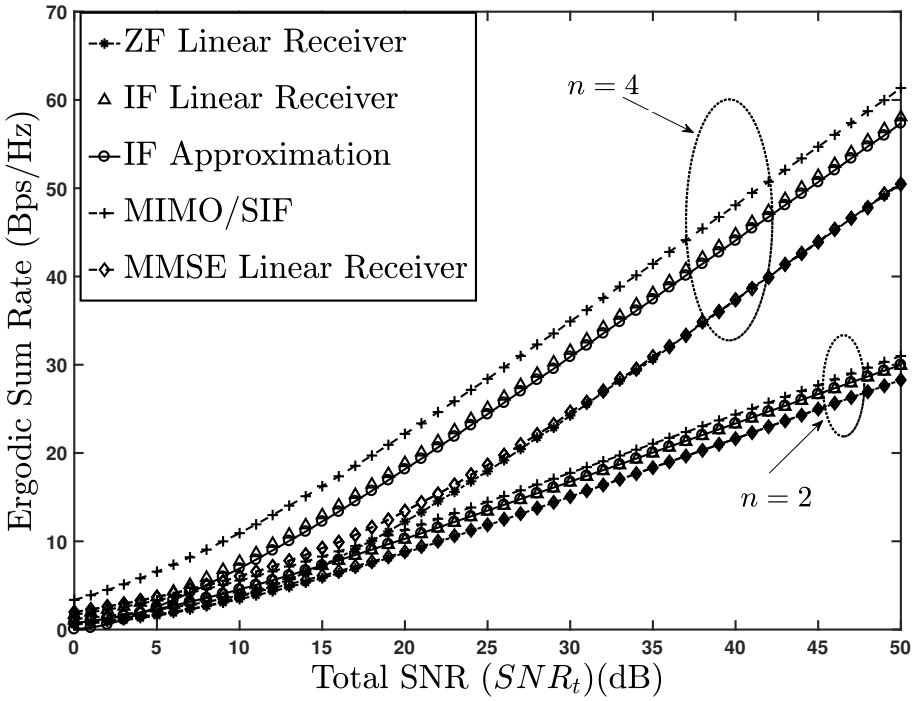

Figure 3.4 – Ergodic Sum Rate versus Total SNR (dB) for the 2 × 2 and 4 × 4 array with uncorrelated channels.

note that the degradation is more severe for the zero forcing linear receiver, showing that IF linear receivers are not only more efficient than ZF linear receiver but also more robust to channel correlation. As the correlation between columns of matrix **H** becomes stronger, the probability of ill-conditioned matrices increases. As ZF linear receivers are significantly

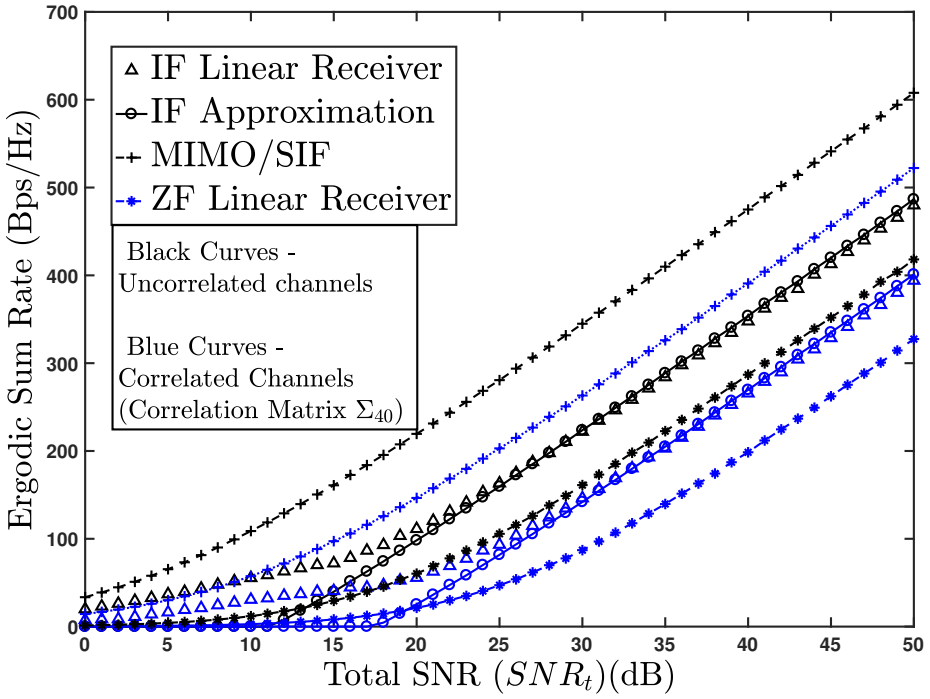

Figure 3.5 – Ergodic Sum Rate versus Total SNR (dB) for the 40×40 array. Black curves and blue curves are for uncorrelated and correlated channels respectively.

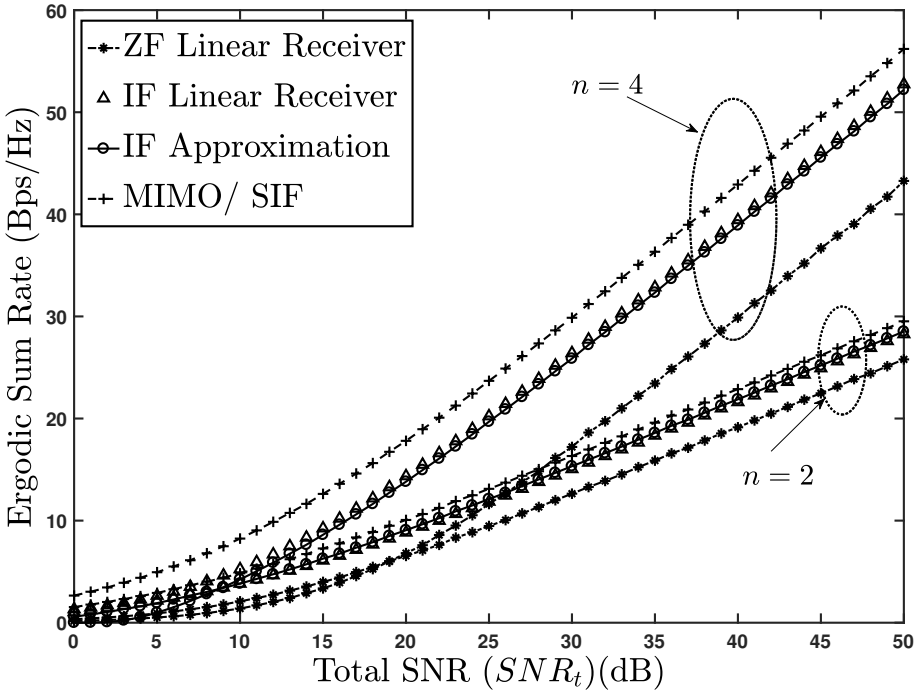

Figure 3.6 – Ergodic Sum Rate versus Total SNR (dB) for the 2×2 array and 4×4 with correlated channels.

suboptimal in these situations, the difference in performance to IF linear becomes clearer.

3.4.4 Outage Probability

In Figures 3.7 we present the outage probability for the 2×2 array for $s = 7dB$, comparing the tightness of both $\hat{R}_{IF_2}$ and $\hat{R}_{IF_n}$ approximations. Both approximations are close to R_{IF} however, $\hat{R}_{IF_2}$ outperforms $\hat{R}_{IF_n}$ which is usually the case for low SNR. For the region of small outage, which is usually the region of interest we can see in Figure 3.8 a more significant difference between the two approximations. It becomes clear that at low SNR, is is better to use the 2×2 array approximation.

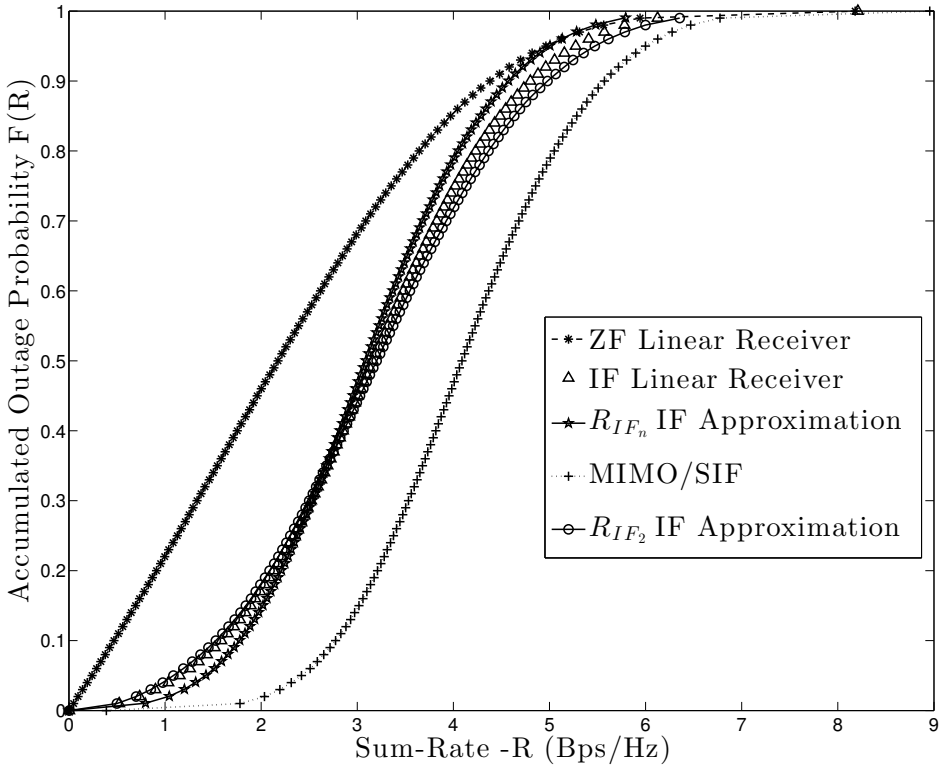

Figure 3.7 – Outage Probability versus Sum Rate for the 2×2 complex array with uncorrelated channels at $s = 7dB$.

3.4.5 Outage Sum Rate

For the same setup as in 3.4.3, we compare the 5% outage sum rate, that is $R_{outage}(0.05)$. Once again, uncorrelated and correlated channels are analyzed.

3.4.5.1 Uncorrelated channels

In Figure 3.9 and 3.10 (black curves) we show the 5% outage rate associated with $\hat{R}_{IF_n}$ for 2×2, 4×4 and 40×40 complex arrays. It is interesting to note that the 5% outage sum rate curves are farther away from each other than the ergodic sum rate curves,

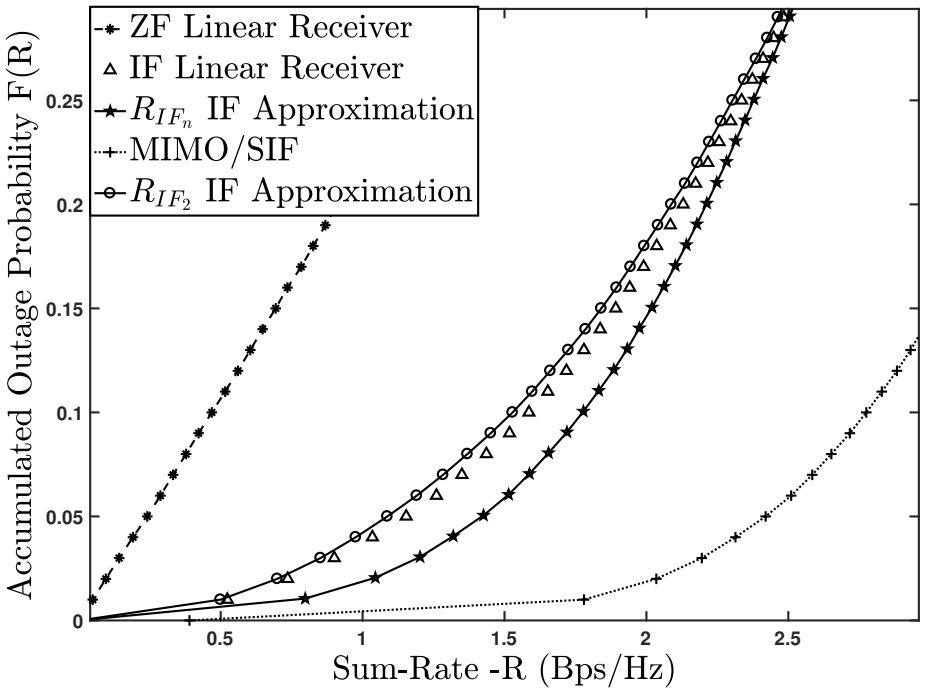

Figure 3.8 – Zoom of the low outage probability region versus Sum Rate for the 2×2 complex array with uncorrelated channels at $s = 7dB$.

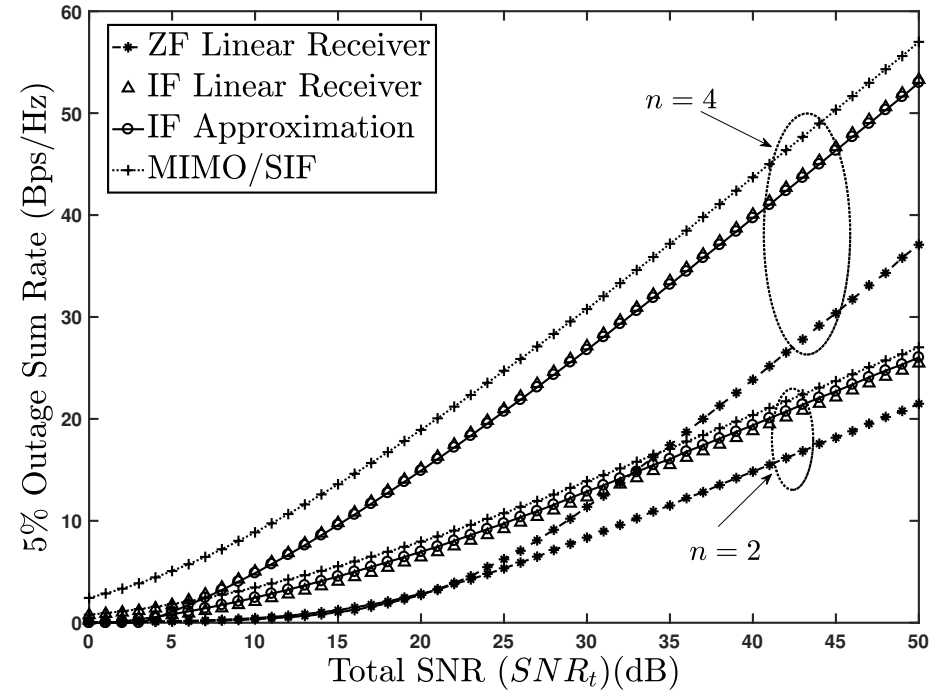

Figure 3.9 – Outage Sum Rate 5% versus Total SNR (dB) for 2×2 and 4×4 arrays with uncorrelated complex channels

showing that for strict outage rate conditions the penalty for using simpler transmission methods is greater.

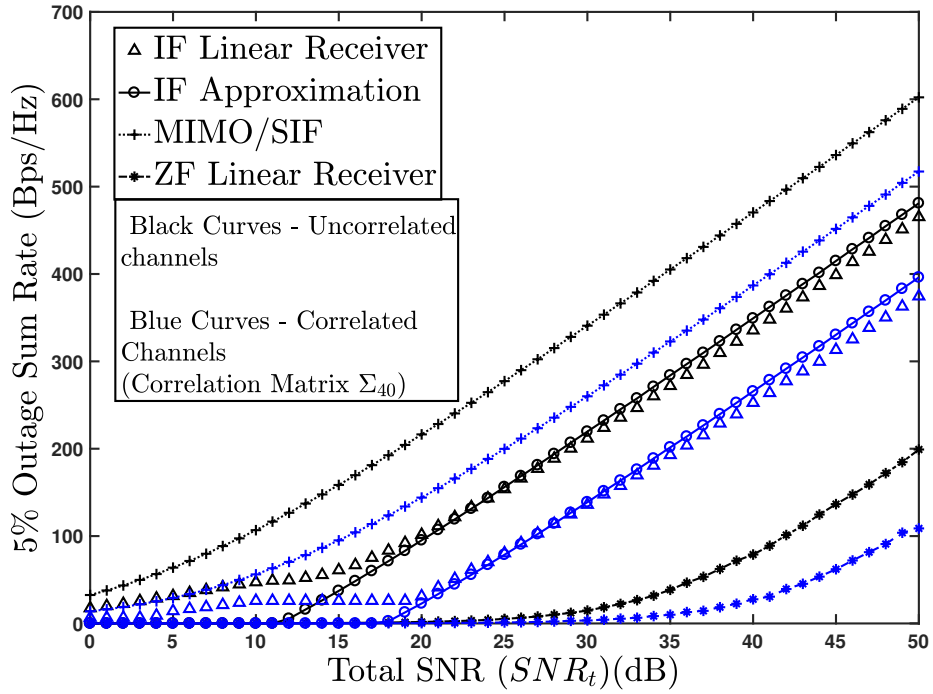

Figure 3.10 – Outage Sum Rate 5% versus Total SNR (dB) for the 40×40 array. Black curves and blue curves are for uncorrelated and correlated channels respectively.

3.4.6 Correlated channels

In Figure 3.11, the outage rate has been plotted using the covariance matrices $\mathbf{\Sigma}_2$ and $\mathbf{\Sigma}_4$. In the same way, the blue curves of Figure 3.10 illustrates the outage rate for the case 40×40 with covariance matrix $\mathbf{\Sigma}_{40}$. In all the cases, a perfect agreement is observed between our proposed approximations and the exact values.

3.5 Chapter conclusions

In this chapter, we have introduced two new approximations to the achievable sum rate of the IF linear receiver in the presence of Rayleigh distributed fading. We have shown, through a series of simulations, that for the 2×2 array, over 98% of channel realizations require only one iteration of the Gauss-Lagrange algorithm to obtain λ_2. We have also shown, that our approximation for the 2×2 array, which is based on the Gauss-Lagrange algorithm, is very tight for all SNR values for both real and complex channels.

We have analytically shown that the $n \times n$ approximation attains the same degrees of freedom as the IF linear receiver and that the coefficient of variation is null for asymptotically high SNR. Through numerical simulations, we have also shown that this approximation is tight for high SNR, and its coefficient of variation is small even for lim-

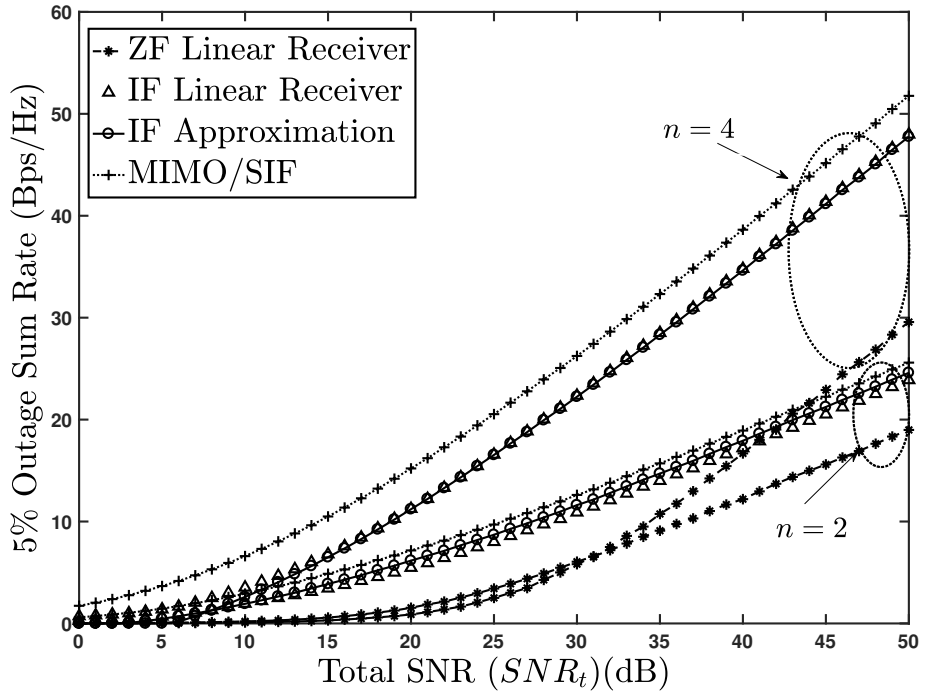

Figure 3.11 – Outage Sum Rate 5% versus Total SNR (dB) for 2 × 2 and 4 × 4 arrays with correlated complex channels.

ited SNR. Observing the relation of the simulations and analytical results, it is possible to state that the approximation is close to the performance of IF linear receivers regardless of channel correlation and number of antennas. From the simulations, it was also possible to observe how the IF linear receiver is more efficient and more robust to channel correlation than the zero forcing linear receiver.

Chapter 4

Approximate Sum Rate for Integer-Forcing Receiver with Imperfect Channel Estimation

4.1 Introduction

In the previous chapter, it is considered that the receiver has perfect channel state information (CSIR). In reality, the channel state has to be estimated and therefore may contain errors. In this chapter, we propose an approximation for the mean achievable sum rate of IF and CFT receivers with imperfect CSI. For IF receivers, we also derive the PDF of the approximation and present the expression of the ergodic IF achievable sum rate for uncorrelated and correlated channels. Only complex channels are taken into consideration. We show the tightness of our approximations through a series of simulations.

The remainder of the chapter is organized as follows. In Section 4.2, we present our approximations and derive its PDF. In Section 4.3, we compare simulation results to those obtained with our approximation. Finally, Section 4.4 concludes the chapter.

4.1.1 Effects of Imperfect CSI

In the traditional framework of the receivers addressed in 2.2, perfect CSI is considered to be available at the receiver. However, in reality, each realization of the channel has to be estimated and therefore it is not perfect. We express the estimated channel gains to be,

$$\mathbf{H}_e = \mathbf{H} + e\mathbf{\Omega}, \tag{4.1}$$

in which, $e\mathbf{\Omega}$ is the estimation error. The matrix $\mathbf{\Omega} \in \mathbb{C}^{n_r \times n_t}$ is uncorrelated to $\mathbf{H}$ and its entries are i.i.d zero mean complex Gaussian with unitary variance. The constant e is a measure of the accuracy of the estimation. In the presence of this kind of channel estimation error, the

Gaussian integer matrix $\mathbf{A}_e$ and the equalization matrix $\mathbf{B}_e$ are calculated respectively as,

$$\mathbf{A}_e = \arg \min_{\substack{\mathbf{A}_e \in Z[i]^{n_t \times n_t} \\ Rank(\mathbf{A}_e)=n_t}} \max_{m=1,\dots,n_t} \|\mathbf{Q}_e \mathbf{a}'_m\|^2, \tag{4.2}$$

$$\mathbf{B}_e = s\mathbf{A}_e\mathbf{H}'_e \left(\mathbf{I}_{n_t} + s\mathbf{H}_e\mathbf{H}'_e\right)^{\dagger}. \tag{4.3}$$

Given the singular value decomposition $\mathbf{V}_e\mathbf{D}_e\mathbf{V}'_e = \mathbf{I}_{n_t} + s\mathbf{H}'_e\mathbf{H}_e$, matrix $\mathbf{Q}_e$ can be calculated as,

$$\mathbf{Q}_e = \mathbf{D}_e^{-1/2}\mathbf{V}'_e. \tag{4.4}$$

It is important to note that matrices $\mathbf{A}_e$ and $\mathbf{B}_e$ might not be optimal for (2.33), therefore resulting in an increase of the effective noise. The achievable sum rates for IF and CFT receivers with imperfect CSI are expressed respectively as,

$$R_{IFe} = n_t \min_m R_{comp,m}\left(\mathbf{H}_e, \mathbf{A}_e, \mathbf{B}_e\right), \tag{4.5}$$

$$R_{CFTe} = \sum_{m=1}^{n_t} R_{comp,m}\left(\mathbf{H}_e, \mathbf{A}_e, \mathbf{B}_e\right), \tag{4.6}$$

4.2 Approximation to the Average Achievable sum rate

In this section, we first derive an approximation to the average achievable sum rate of both IF and CFT receivers. Next, we analyze the rate-loss associated with the imperfect channel estimation. In the third subsection, we derive an approximation to the IF sum rate PDF for uncorrelated and correlated random channel matrices. In the last subsection, we define and present an approximation for the ergodic IF sum rate.

4.2.1 Average Achievable sum rate of IF and CFT receivers

We start by defining our parameter of analysis.

Definition 4.2.1. Considering the channel matrix $\mathbf{H}$ to be fixed and the other system parameters to be distributed according to 2.2, the average achievable sum rate is defined as,

$$\bar{R}_F = \mathbb{E}_{\mathbf{\Omega},\mathbf{Z},\mathbf{X}}\left[R_F\left(\mathbf{\Omega}, \mathbf{Z}, \mathbf{X}\right)\right], \tag{4.7}$$

in which R_F is a general notation representing either R_{IF_e} or R_{CFT_e}.

In order to obtain the expectation of the achievable sum rates, it is simpler to first obtain the expectation of the effective noise variance for the m-th stream and then calculate $\bar{R}_F$.

We begin the analysis by obtaining an approximation to the inverse of the estimated channel matrix $\mathbf{H}_e$. Considering $e << 1$, the linear part of a Taylor series is used to approximate $\mathbf{H}_e$ as,

$$\mathbf{H}_e^{\dagger} = (\mathbf{H} + e\mathbf{\Omega})^{\dagger} \approx \mathbf{H}^{\dagger}\left(\mathbf{I}_n - e\mathbf{\Omega}\mathbf{H}^{\dagger}\right). \tag{4.8}$$

At high SNR, $\mathbf{B}_e$ is approximated as $\mathbf{B}_e \approx \mathbf{A}_e\mathbf{H}_e^{\dagger}$, and using (4.8), it is possible to obtain,

$$\mathbf{B}_e \approx \mathbf{A}_e\mathbf{H}^{\dagger}\left(\mathbf{I}_n - e\mathbf{\Omega}\mathbf{H}^{\dagger}\right). \tag{4.9}$$

The effective noise under imperfect CSI is approximated as,

$$\begin{aligned}\mathbf{Z}_{ef} \approx& (\mathbf{B}_e\mathbf{H} - \mathbf{A}_e)\mathbf{X} + \mathbf{B}_e\mathbf{Z} \\ \approx& \mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{Z} - e\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{X} - e\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{H}^{\dagger}\mathbf{Z}.\end{aligned} \tag{4.10}$$

The effective noise variance of the m-th data stream in the presence of imperfect CSI is calculated as $\sigma_{e,m}^2 = \mathbb{E}\left[\mathbf{z}_{ef,m}\mathbf{z}_{ef,m}'\right]_{\mathbf{Z},\mathbf{X},\mathbf{\Omega}}$. Here, $\mathbf{z}_{ef,m}'$ is the m-th row of $\mathbf{Z}_{ef}$. The equation for the squared effective noise is,

$$\begin{aligned}\mathbf{Z}_{ef}\mathbf{Z}_{ef}' =& \mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{Z}\mathbf{Z}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' - e\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{Z}\mathbf{X}'\mathbf{\Omega}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' - e\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{X}\mathbf{Z}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' \\ &+ e^2\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{X}\mathbf{X}'\mathbf{\Omega}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' - e\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{Z}\mathbf{Z}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{\Omega}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' - e\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{H}^{\dagger}\mathbf{Z}\mathbf{Z}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' \\ &+ e^2\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{X}\mathbf{Z}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{\Omega}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' + e^2\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{H}^{\dagger}\mathbf{Z}\mathbf{X}'\mathbf{\Omega}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' \\ &+ e^2\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{H}^{\dagger}\mathbf{Z}\mathbf{Z}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{\Omega}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e'.\end{aligned} \tag{4.11}$$

We first calculate the expected value with respect to $\mathbf{Z}$ and $\mathbf{X}$ using the following identities,

$$\mathbb{E}_{\mathbf{X}}[\mathbf{X}] = \mathbf{0}_{n\times l}, \tag{4.12}$$

$$\mathbb{E}_{\mathbf{X}}\left[\mathbf{X}\mathbf{X}'\right] = s\mathbf{I}_n, \tag{4.13}$$

$$\mathbb{E}_{\mathbf{Z}}[\mathbf{Z}] = \mathbf{0}_{q\times l} \tag{4.14}$$

$$\mathbb{E}_{\mathbf{Z}}\left[\mathbf{Z}\mathbf{Z}'\right] = N_0\mathbf{I}_q. \tag{4.15}$$

Here $\mathbf{0}_{a\times b}$ represents the matrix with all entries equal to 0 with a rows and b columns. The result of this calculation is given as,

$$\begin{aligned}\mathbb{E}\left[\mathbf{Z}_{ef}\mathbf{Z}_{ef}'\right]_{\mathbf{Z},\mathbf{X}} =& N_0\mathbf{A}_e\mathbf{H}^{\dagger}\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' + e^2s\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{\Omega}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' - eN_0\mathbf{A}_e\mathbf{H}^{\dagger}\left(\mathbf{H}'\right)^{\dagger}\mathbf{\Omega}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' \\ &+ e^2N_0\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{H}^{\dagger}\left(\mathbf{H}'\right)^{\dagger}\mathbf{\Omega}'\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e' - eN_0\mathbf{A}_e\mathbf{H}^{\dagger}\mathbf{\Omega}\mathbf{H}^{\dagger}\left(\mathbf{H}'\right)^{\dagger}\mathbf{A}_e'.\end{aligned} \tag{4.16}$$

Here, N_0 represents the noise spectral density.

4.2.1.1 Relation of $\mathbf{A}_e$ and Ω

According to (4.2) and (4.4) it is possible to observe that,

$$\mathbf{A}_e = \arg \min_{\substack{\mathbf{A}_e\in Z[i]^{n\times n} \\ Rank(\mathbf{A}_e)=n}} \max_{m=1,\ldots,n} \mathbf{a}_{e,m}'\left(\mathbf{I}_n + s\left(\mathbf{H} + e\mathbf{\Omega}\right)'\left(\mathbf{H} + e\mathbf{\Omega}\right)\right)^{-1}\mathbf{a}_{e,m}. \tag{4.17}$$

Here $\mathbf{a}_{e,m}$ represents the m-th row of $\mathbf{A}_e$. As can be seen in (4.17), matrix $\mathbf{A}_e$ is a function of $\mathbf{\Omega}$, and therefore, it is complex to obtain the expected value of (4.16) with respect to $\mathbf{\Omega}$. However, if e is small the dependency between these two random matrices can be considered negligible for the calculation of σ_e^2. In the next paragraphs, we present the motivations for this assumption.

At high SNR, the optimization (4.17) can be simplified to,

$$\mathbf{A}_e = \arg \min_{\substack{\mathbf{A}_e \in Z[i]^{n\times n} \\ Rank(\mathbf{A}_e)=n}} \max_{m=1,\ldots,n} \| \left((\mathbf{H} + e\mathbf{\Omega})' \right)^{\dagger} \mathbf{a}_{e,m} \|^2. \tag{4.18}$$

As stated in section II.B, the optimization (4.18) can be viewed as a SIVP of a lattice with basis $\mathbf{G} = \left((\mathbf{H} + e\mathbf{\Omega})' \right)^{\dagger}$. Analogously, for perfect CSI and high SNR the optimization (4.2) can be performed by solving a SIVP of $\Lambda((\mathbf{H}')^{\dagger})$.

From (4.18) it is possible to observe that as the quantity e decreases, lattice $\Lambda(\mathbf{G})$ approximates $\Lambda\left((\mathbf{H}')^{\dagger}\right)$, and thus, the points of both lattices converge. Therefore solving the SIVP for $\Lambda(\mathbf{G})$ in this scenario, likely results in an unimodular matrix $\mathbf{A}$ which is also optimal for $\Lambda\left((\mathbf{H}')^{\dagger}\right)$. In Figure 4.1 we present an example to illustrate this effect using the following parameters,

$$\mathbf{H} = \begin{bmatrix} 1.80 & -1.36 \\ -2.71 & 3.88 \end{bmatrix} \tag{4.19}$$

$$\mathbf{H}_e = \begin{bmatrix} 1.63 & -1.24 \\ -2.71 & 3.81 \end{bmatrix}, \tag{4.20}$$

and $e = 0.2$. For both lattices the optimal integer matrix is,

$$\mathbf{A} = \begin{bmatrix} -1 & -2 \\ 1 & 3 \end{bmatrix}. \tag{4.21}$$

This is the reason for the low dependency between matrices $\mathbf{\Omega}$ and $\mathbf{A}_e$ when e is small. In section 4.3 we present simulations that measure the dependency between the elements of $\mathbf{\Omega}$ and $\mathbf{A}_e$ for $e = 0.2$ and various antenna arrays.

4.2.1.2 Average achievable sum rate approximation

We continue the derivation of the approximations by applying the following identities to 4.16,

$$\mathbb{E}_{\mathbf{\Omega}}[\mathbf{\Omega}] = \mathbf{0}_{q\times n}, \tag{4.22}$$

$$\mathbb{E}_{\mathbf{\Omega}}\left[\mathbf{\Omega}\mathbf{\Omega}'\right] = n\mathbf{I}_q, \tag{4.23}$$

$$\mathbb{E}_{\mathbf{\Omega}}\left[\mathbf{\Omega}\left(\mathbf{H}'\mathbf{H}\right)^{\dagger}\mathbf{\Omega}'\right] = \mathrm{Tr}\left(\left(\mathbf{H}'\mathbf{H}\right)^{\dagger}\right)\mathbf{I}_q, \tag{4.24}$$

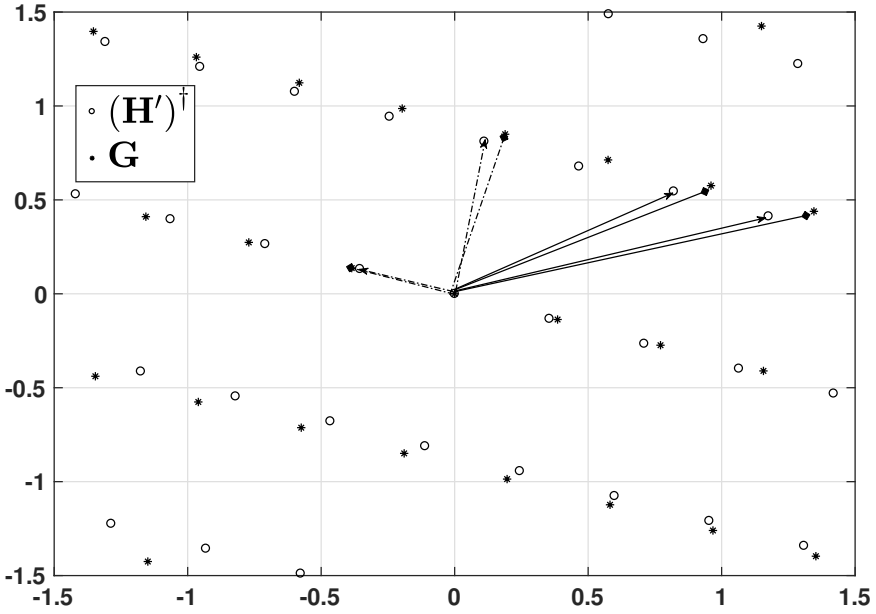

Figure 4.1 – Lattice points for $\Lambda((\mathbf{H}')^{\dagger})$ and $\Lambda(\mathbf{G})$. The solid arrows represent the vector basis before lattice reduction and the dotted arrows present the reduced vector basis.

ignoring the dependency of $\mathbf{A}_e$ and $\mathbf{\Omega}$ we obtain,

$$\sigma^2_{e,m} = \mathbb{E}\left[\mathbf{z}_{ef,m}\mathbf{z}'_{ef,m}\right]_{\mathbf{Z},\mathbf{X},\mathbf{\Omega}} \approx \|\left(\mathbf{H}'\right)^{\dagger}\mathbf{a}'_{e,m}\|^2\left(N_0 + se^2n + N_0e^2\mathrm{Tr}\left(\left(\mathbf{H}'\mathbf{H}\right)^{\dagger}\right)\right), \tag{4.25}$$

As we are analyzing the system at high SNR, it is possible to observe that,

$$\|\left(\mathbf{H}'\right)^{\dagger}\mathbf{a}'_{e,m}\|^2 \approx s\|\mathbf{Q}\mathbf{a}'_{e,m}\|^2 \approx s\lambda^2_m(\mathbf{Q}_e) \tag{4.26}$$

Therefore the effective noise variance of the m-th stream is approximated as

$$\sigma^2_{e,m} \approx s\lambda^{-2}_m(\mathbf{Q}_e)\left(N_0 + se^2n + N_0e^2\mathrm{Tr}\left(\left(\mathbf{H}'\mathbf{H}\right)^{\dagger}\right)\right). \tag{4.27}$$

Finally the approximations to the average achievable sum rate of IF and CFT receivers in the presence of imperfect CSI, $\bar{R}_{IFe}$ and $\bar{R}_{CFTe}$ respectively, are expressed as

$$\bar{R}_{IFe} = n\log^+\left(\lambda^{-2}_n(\mathbf{Q}_e)\left(N_0 + se^2n + N_0e^2\mathrm{Tr}\left(\left(\mathbf{H}'\mathbf{H}\right)^{\dagger}\right)\right)^{-1}\right), \tag{4.28}$$

$$\bar{R}_{CFTe} = \sum_{i=1}^{n}\log^+\left(\lambda^{-2}_i(\mathbf{Q}_e)\left(N_0 + se^2n + N_0e^2\mathrm{Tr}\left(\left(\mathbf{H}'\mathbf{H}\right)^{\dagger}\right)\right)^{-1}\right). \tag{4.29}$$

4.2.2 Distribution of the Approximation of the Achievable IF sum rate under Imperfect CSI

In this subsection we derive the distribution of an approximation of $\bar{R}_{IFe}$ and, with this result, we also derive an approximation to the ergodic IF sum rate. We assume that

the elements of the channel matrix $\mathbf{H}$ are complex Gaussian distributed random variables (with Rayleigh distributed envelope). Concerning channel correlation, we consider the two scenarios presented in 2.2. For both scenarios, we are able to derive an approximation to the PDF of $\bar{R}_{IFe}$ for a system with $n_r = n_t = n$ antennas. This restriction is necessary as we use the approximation derived in the previous chapter. The largest successive minima of a lattice with basis $\mathbf{Q}_e$ can be approximated as,

$$\lambda_n(\mathbf{Q}_e) \approx \sqrt{\gamma_n}\det(\mathbf{Q}_e'\mathbf{Q}_e)^{1/2n}. \tag{4.30}$$

Therefore $\bar{R}_{IFe}$ can be approximated as,

$$\bar{R}_{IFe} \approx R_{IF_{ap}} = \max\Big(0, -n\log\Big(\gamma_n \det(\mathbf{Q}_e'\mathbf{Q}_e)^{1/n}\Big(N_0 + se^2n + N_0e^2\mathrm{Tr}\Big(\big(\mathbf{H}'\mathbf{H}\big)^{\dagger}\Big)\Big)\Big)\Big). \tag{4.31}$$

Since $\mathbf{Q}_e'\mathbf{Q}_e = \mathbf{I}_n + s\mathbf{H}_e'\mathbf{H}_e$, it is possible to express $R_{IF_{ap}}$ as,

$$R_{IF_{ap}} = \max\Big(0, -\log(\gamma_n^n) + C_{\mathrm{MIMO}_e} - \log\Big(\Big(N_0 + se^2n + N_0e^2\mathrm{Tr}\Big(\big(\mathbf{H}'\mathbf{H}\big)^{\dagger}\Big)\Big)^n\Big)\Big), \tag{4.32}$$

where we define C_{MIMO_e} as,

$$C_{\mathrm{MIMO}_e} = \log\Big(\det\big(\mathbf{I}_n + s\mathbf{H}_e'\mathbf{H}_e\big)\Big). \tag{4.33}$$

The covariance matrix of $\mathbf{H}_e$ is calculated as,

$$\mathbf{\Sigma}_{e,i,j} = \begin{cases} 1 + e^2, & \text{if } i = j \\ \mathbf{\Sigma}_{i,j}, & \text{otherwise.} \end{cases} \tag{4.34}$$

Without loss of generality the covariance matrix is normalized to have unit diagonal elements, and thus the constant $s_e = s(1+e^2)$ and covariance matrix $\mathbf{\Sigma}_f = \mathbf{\Sigma}_e/(1+e^2)$ are used to describe the correlated distributions of C_{MIMO_e}. The approximation to the average achievable IF sum rate is expressed as the classical MIMO capacity, with uniform power allocation [39], plus a constant. The distribution of (4.33) is given as [40],

$$p_{\mathrm{MIMO}_e}(r) = \frac{1}{2\pi}\int_R e^{\iota\omega r}\Phi_e(\omega)d\omega. \tag{4.35}$$

Here $\Phi(\omega)$ represents the appropriate characteristic function. We now divide our analysis into the correlated and uncorrelated scenarios.

4.2.2.1 Uncorrelated channels

For uncorrelated channels the characteristic function is given as [40],

$$\Phi_{un,e}(\omega) = K\det(hu_{j,k})_{j,k=1,\ldots,n}, \tag{4.36}$$

with

$$K^{\dagger} = \frac{1}{n!}\prod_{j=1}^{n}\Gamma(j+1)\Gamma(j+2n), \tag{4.37}$$

$$hu_{j,k} = s_e^{-\imath\omega}\Gamma(j+k-1)U(\imath\omega, \imath\omega - j - k + 2; s_e^{-1}), \tag{4.38}$$

Furthermore according to [25], for practical number of antennas, the quantity

$$T = N_0 e^2 \mathrm{Tr}\left(\left(\mathbf{H}'\mathbf{H}\right)^{\dagger}\right), \tag{4.39}$$

is considered to be small, and therefore can be neglected.

4.2.2.2 Correlated channels

Considering the Kronecker model with only transmit-side spatial correlation, the channel matrix can be expressed as,

$$\mathbf{H} = \mathbf{H}_w \boldsymbol{\Sigma}'^{1/2}. \tag{4.40}$$

The matrix $\mathbf{H}_w$ is considered to be uncorrelated Rayleigh distributed and is post-multiplied by the transmit-side spatial correlation matrix $\boldsymbol{\Sigma}$. Analyzing the quantity T, we have the following,

$$\begin{aligned} T =& N_0 e^2 \mathrm{Tr}\left(\left(\mathbf{H}'\mathbf{H}\right)^{\dagger}\right) && (4.41)\\ =& N_0 e^2 \mathrm{Tr}\left(\left(\boldsymbol{\Sigma}^{1/2}\mathbf{H}_w'\mathbf{H}_w\boldsymbol{\Sigma}'^{1/2}\right)^{\dagger}\right) && (4.42)\\ =& N_0 e^2 \sum_{i=1}^{n}\rho_i^{-1}\left(\boldsymbol{\Sigma}^{1/2}\mathbf{H}_w'\mathbf{H}_w\boldsymbol{\Sigma}'^{1/2}\right) && (4.43)\\ \leq& N_0 e^2 \sum_{i=1}^{n}\rho_i^{-1}\left(\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}'^{1/2}\right)\rho_i^{-1}\left(\mathbf{H}_w\mathbf{H}_w'\right) && (4.44)\\ \leq& n\rho_{\min}^{-1}\left(\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}'^{1/2}\right)N_0 e^2 \mathrm{Tr}\left(\left(\mathbf{H}_w'\mathbf{H}_w\right)^{\dagger}\right). && (4.45)\end{aligned}$$

Here the operators $\rho_i(\mathbf{M})$ and $\rho_{\min}(\mathbf{M})$ represent the i-th and the minimum eigenvalues of matrix $\mathbf{M}$, respectively. Since $N_0 e^2 \mathrm{Tr}\left(\left(\mathbf{H}_w'\mathbf{H}_w\right)^{\dagger}\right)$ is known to be small [25] for a practical number of antennas, it is possible to observe from (4.45) that if $n\rho_{\min}^{-1}\left(\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}'^{1/2}\right)$ is not too large, the quantity T can be neglected. Throughout the chapter, only channels with small T will be considered due to the complexity of a more general analysis.

The characteristic function of (4.33) for correlated channels is expressed as [40],

$$\Phi_{cor,e}(\omega) = K_{\boldsymbol{\Sigma}_f}\det(hc_{j,k})_{j,k=1,\ldots,n}, \tag{4.46}$$

in which,

$$K_{\mathbf{\Sigma}_f} = \frac{-n!}{\Gamma(n+1)\prod_{j>k}(\sigma_j - \sigma_k)} \prod_{j=1}^{n} \frac{\rho_j^{-n}\left(\mathbf{\Sigma}_f\right)}{\Gamma(j)} \tag{4.47}$$

$$hc_{j,k} = \rho_j^{k-\imath\omega}\left(\mathbf{\Sigma}_f\right) s_e^{-\imath\omega}\Gamma(k) U\left(\imath\omega, \imath\omega - k + 1; s_e^{-1}\rho_j^{-1}\left(\mathbf{\Sigma}_f\right)\right). \tag{4.48}$$

4.2.2.3 Distribution of $R_{IF_{ap}}$

Now, since $R_{IF_{ap}}$ is a linear function of C_{MIMO}, the PDF of the approximation can be expressed as,

$$p_{R_{IFap}}(r) = \begin{cases} 0, & \text{if } r < 0 \\ K_n \delta(r) + p_{\text{MIMO}_e}\left(r + \log(\gamma_n^n (N_0 + s_e e^2 n)^n)\right), & \text{otherwise} \end{cases} \tag{4.49}$$

in which $\delta(.)$ represents the dirac delta and K_n is given by,

$$K_n = \int_{-\infty}^{0} p_{\text{MIMO}_e}\left(r + \log(\gamma_n^n (N_0 + s_e e^2 n)^n)\right) dr. \tag{4.50}$$

It is important to note that p_{MIMO_e} is obtained from the appropriate characteristic function ($\Phi_{un,e}$ for uncorrelated and $\Phi_{cor,e}$ for correlated channels).

4.2.3 Ergodic sum rate of IF receivers with imperfect CSI

In this subsection, we define and provide an approximation to the the ergodic IF sum rate.

Definition 4.2.2. The ergodic sum rate is the average transmission rate with respect to the channel realizations as,

$$R_{IFa} = \mathbb{E}_{\mathbf{H}}\left[\bar{R}_{IF}\right] \tag{4.51}$$

Using (4.49), we approximate R_{IFa} as the mean value of $R_{IF_{ap}}$ with respect to $\mathbf{H}$, which can be obtained as,

$$R_{IFa} = \mathbb{E}\left[R_{IF_{ap}}\right]_{\mathbf{H}} = \int_{-\infty}^{\infty} r p_{R_{IF_{ap}}}(r) dr. \tag{4.52}$$

However, when the characteristic function of C_{MIMO} exists around 0, it is simpler to obtain the mean value $R_{IF_{ap}}$ as a function of the derivative of $\Phi(\omega)$,

$$R_{IFa} = \log\left(\gamma_n^n (N_0 + s e^2 n)^n\right) - \imath \frac{d\Phi_e(\omega)}{d\omega}\bigg|_{\omega=0}. \tag{4.53}$$

4.3 Simulations

In this section, we present a series of simulations in order to show the accuracy of our approximations. For all simulations it is considered that the system variables are distributed according to the system model described in 2.2 and 4.1.1.

4.3.1 $\mathbf{A}_e$ and Ω dependency

In this subsection, we analyze the dependency between $\mathbf{A}_e$ and $\mathbf{\Omega}$ through a series of simulations. The parameters used in all simulations are the same, channel gains are distributed as described in 2.2, 4.1.1 and 10^4 repetitions of a Monte-Carlo simulation were performed for each antenna array. Arrays with n receiving and transmitting antennas ($n \times n$) were used, with $n = 2 \ldots 15$ and quantity $e = 0.2$. In Figure 4.2 the maximum Pearson correlation coefficient [42] between all pairs of elements from matrices $\mathbf{A}_e$ and $\mathbf{\Omega}$ were calculated. It is possible to see that, as the number of antennas grows, the maximum correlation coefficient saturates at 0.015, indicating a very small correlation between the variables of each matrix.

Although widely used, the Pearson correlation coefficient is unable to indicate

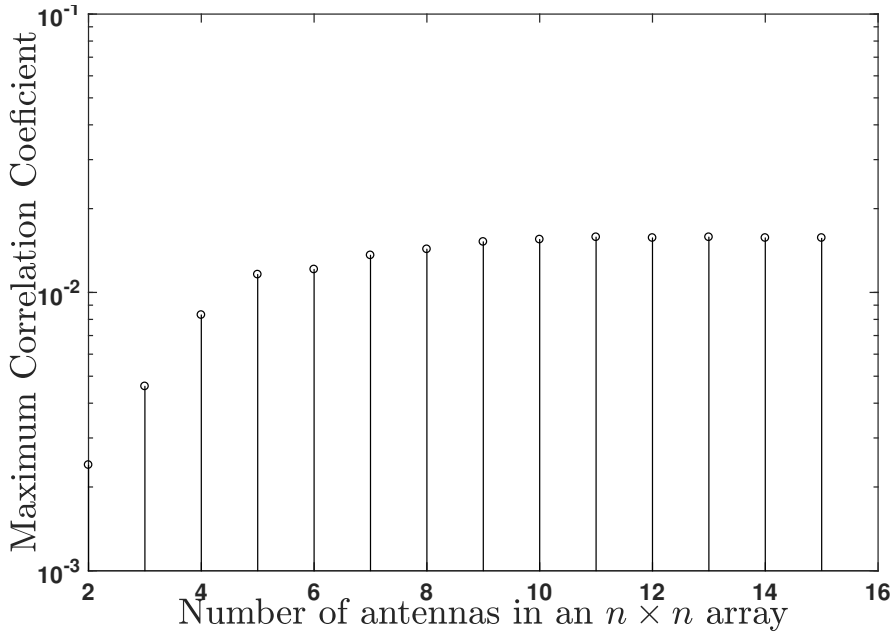

Figure 4.2 – Number of antennas in each array vs Maximum correlation coefficient between elements of $\mathbf{\Omega}$ and $\mathbf{A}_e$

.

independence between variables. Therefore in our second simulation, we analyzed the bias-corrected distance correlation [43]. It is known that if the distance correlation is null, the two

random variables are independent. The implementation of the algorithm described in [43] allows us to derive the distance correlation between random vectors and also performs a t-test of independence. The t-test provides the p-value, a measure of how likely it is to obtain a distance correlation equal to or larger than the test result if the experiment is repeated, assuming that the variables are indeed independent. The maximum distance correlation obtained through our simulations was 1.38×10^{-4} and the minimum p-value obtained was 0.16, indicating that there is no significant statistical dependency between the realizations of matrices $\mathbf{\Omega}$ and $\mathbf{A}_e$ for small e.

4.3.2 Average achievable IF and CFT sum rate

In this subsection, we compare the simulated average achievable sum rates of the IF and CFT receivers with their proposed approximations (4.28) and (4.29), in Fig. 4.3 and Fig. 4.4, respectively. The channel matrix

$$\mathbf{H}_s = \begin{pmatrix} 0.9+0.9i & 1.1+1.1i \\ 0.6+0.6i & 1.4+1.4i \end{pmatrix} \tag{4.54}$$

is considered to be fixed while other system variables are distributed according to the system model described in 2.2 and 4.1.1. This specific channel matrix was chosen as it has a large condition number and under this condition IF receivers significantly outperform simpler linear receivers such as ZF and MMSE. For all scenarios, the corresponding curves for $\bar{R}_{IF_e}$

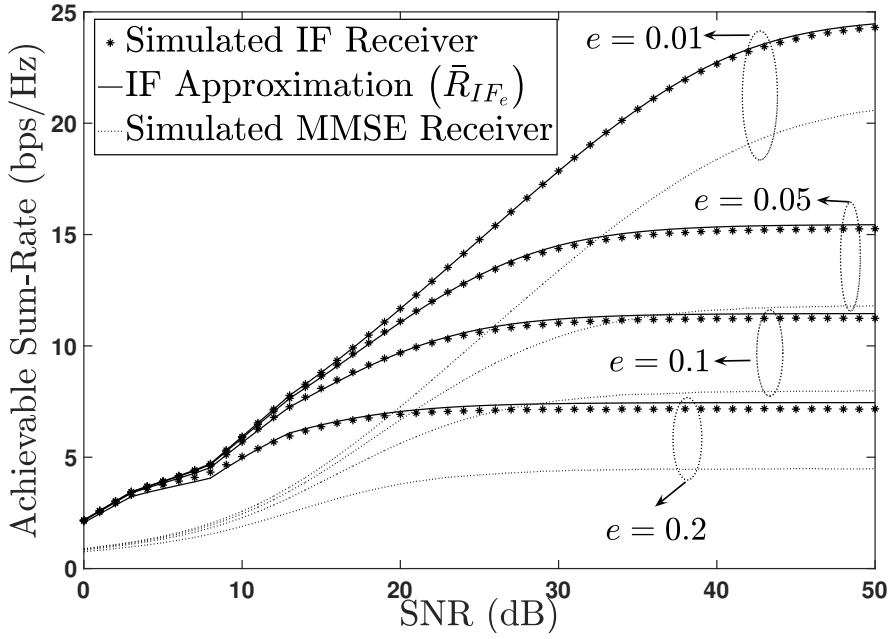

Figure 4.3 – SNR vs Mean achievable IF sum rate $\left(\bar{R}_{IF_e}\right)$ for channel matrix $\mathbf{H}_s$.

and $\bar{R}_{CFT_e}$ are very close to the simulations, although the difference becomes larger as the

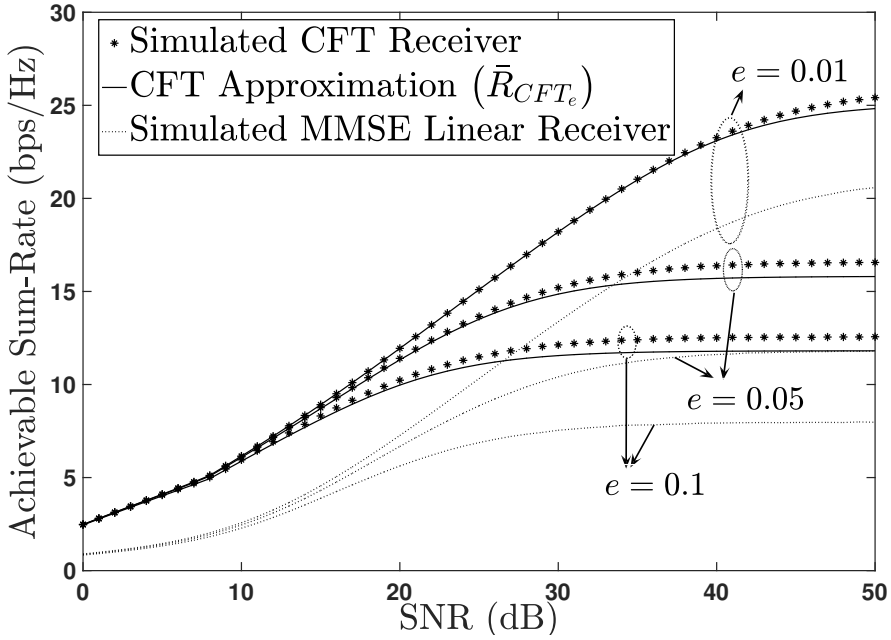

Figure 4.4 – SNR vs Mean achievable CFT sum rate $\left(\bar{R}_{CFT_e}\right)$ for channel matrix $\mathbf{H}_s$.

value of e increases. It is interesting to note that for all cases, MMSE receivers performed worse than IF and CFT receivers.

4.3.3 Ergodic IF achievable sum rate

In this subsection we consider that the elements of channel matrix $\mathbf{H}$ are distributed as described in 4.2.2. We compare curves for the simulated ergodic achievable IF and MMSE sum rates with the approximation R_{IFa}. In Figure 4.5 and 4.6 we variate the quantity e and the number of antennas, respectively, in the array for the uncorrelated channel scenario. For the correlated scenario, presented in Figures 4.7 and 4.8 all covariance matrices follow the same construction method, in which all elements are equal to 0.7 with the exception of the elements in the diagonal, which are unit. Once again we variate the quantity e and the number of antennas in the arrays. It is possible to observe that for all cases, the simulation is very close to our approximation at high SNR. It is also interesting to see that the difference in performance between the IF and MMSE receivers greatly increases as correlation becomes larger.

4.3.4 Variable e

So far, we considered e as a constant. However, in more realistic models it is considered to be a function of several other system parameters and variables. In this subsection, we use one of the models presented in [44] which considers the channel as block fading and

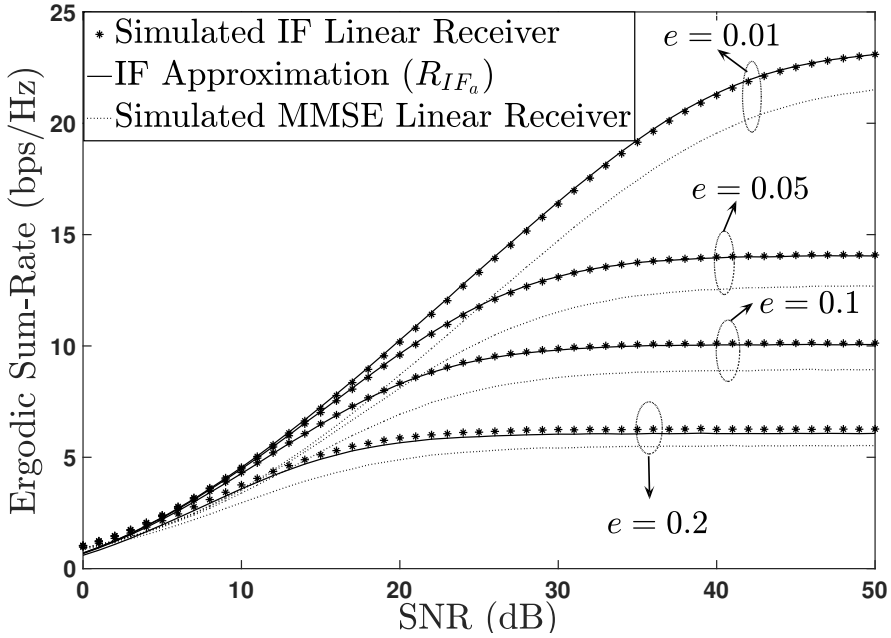

Figure 4.5 – Signal to noise ratio versus Ergodic achievable IF sum rate for uncorrelated channels

.

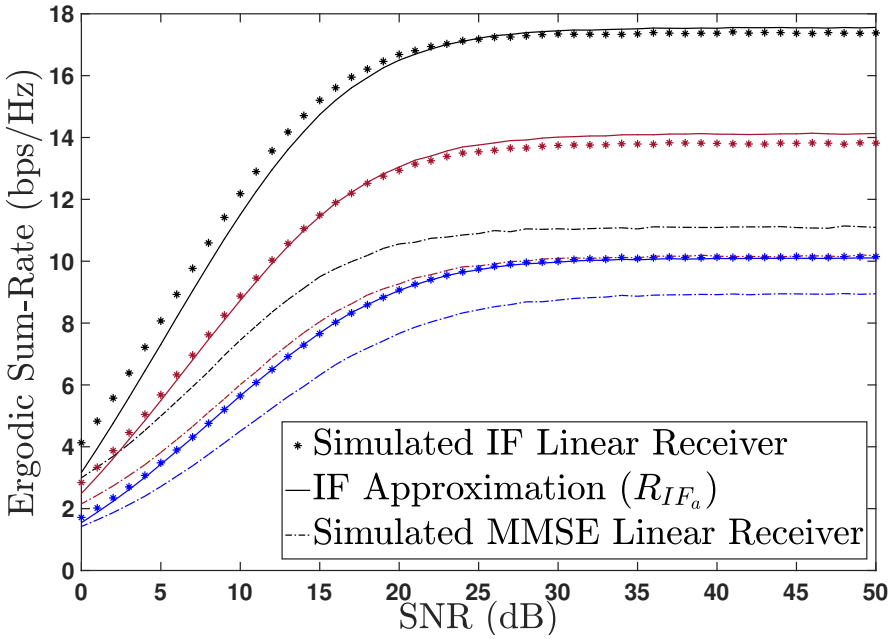

Figure 4.6 – Signal to noise ratio versus Ergodic achievable IF sum rate for uncorrelated channels. Black curves are for $n = 4$, red for $n = 3$ and blue for $n = 2$ in a $n \times n$ array. The quantity $e = 0.1$ is used for all curves.

.

the estimation is implemented via averaging over L noisy pilot symbols. Moreover, we consider the scenario where the power allocated to pilot symbols (s_p), is not necessarily equal to the power allocated to data transmission. If perfect orthogonality between the pilot se-

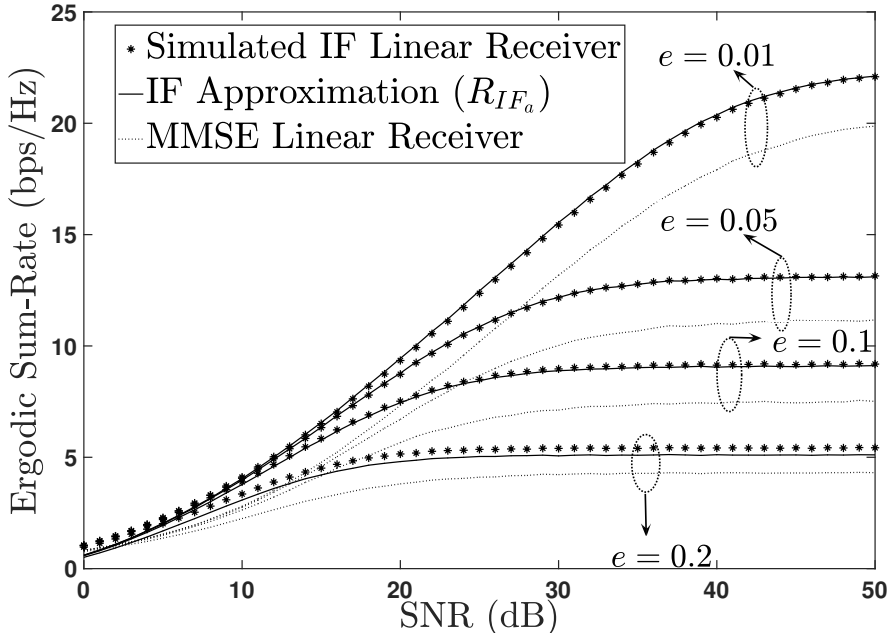

Figure 4.7 – Signal to noise ratio versus Ergodic achievable IF sum rate for correlated channels

.

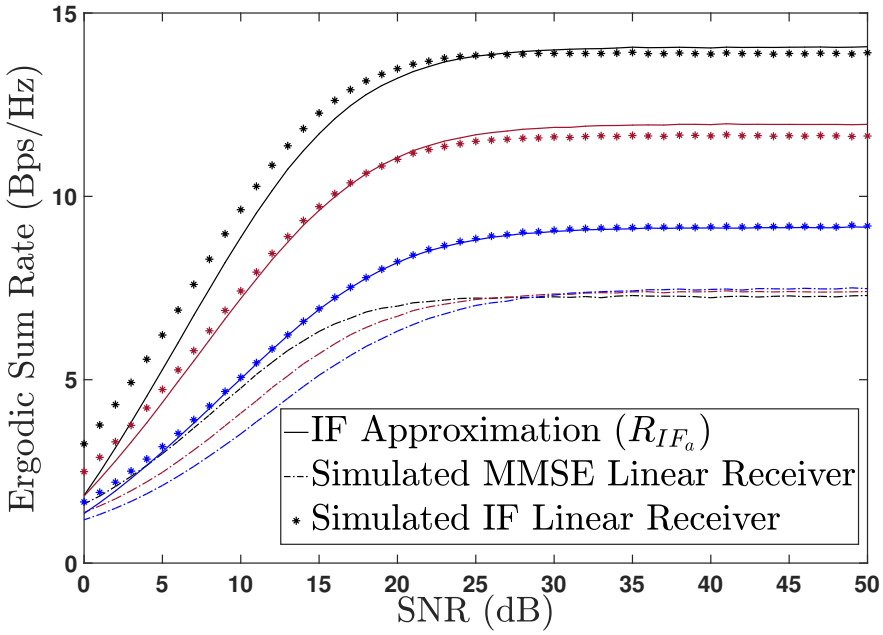

Figure 4.8 – Signal to noise ratio versus Ergodic achievable IF sum rate for correlated channels. Black curves are for $n = 4$, red for $n = 3$ and blue for $n = 2$ in a $n \times n$ array. The quantity $e = 0.1$ is used for all curves

.

quences is assumed, e is calculated as,

$$e = \frac{s^{-4(\alpha+1)}\left(2\beta L s^{2(\alpha+1)} + 1\right)}{\beta^2 L^2}. \tag{4.55}$$

Here α and β are system parameters which relate signal to noise ratio of the pilots and data symbols as $s_p = \beta s^{-\alpha} s$. In Figure 4.9, we compare the simulated ergodic IF and MMSE sum rate with our approximation with e calculated as (4.55), for various $n \times n$ antenna arrays. We used $L = n$, as this represents the smallest possible pilot length size that maintains orthogonality among pilot sequences. It is interesting to note that since the value of e decreases as

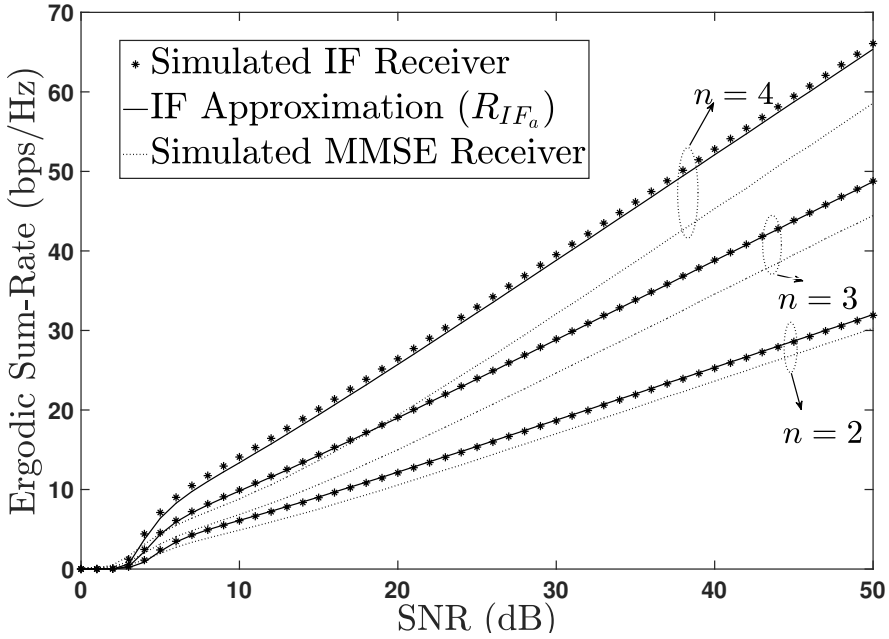

Figure 4.9 – Signal to noise ratio versus Ergodic achievable IF sum rate for uncorrelated channels with variable e. Parameters $\alpha = 0.2$, $\beta = 0.2$ and $L = n$.

.

the s becomes larger, there is no saturation in the transmitted sum rate.

4.4 Chapter Conclusion

In this chapter, we have presented an approximation to the IF and CFT receivers average sum rate in the presence of channel estimation error for fixed channel realizations. We have also derived an approximation to the Ergodic IF sum rate for correlated and uncorrelated channels. Through a series of simulations, it was shown that the approximations presented are tight for all SNR values and small channel estimation error coefficient e values. Most important is the fact that even for small values of e, the system suffers a serious degradation of its sum rate.

Chapter 5

Phase Corrected ICA for Uplink Massive MIMO

5.1 Introduction

In this chapter, we work with the uplink of a Massive MIMO system. We propose a modified version of the fastICA algorithm called Phase Corrected ICA (PCICA). Exploring the format of commonly used constellations, it increases the performance of blind decoders based on ICA by iteratively selecting vectors that lead to estimated constellations that are more aligned with the axis. In section II we introduce a slightly different system model, more suitable to describe a multi-cell scenario. In section III fastICA is reviewed and section IV introduces the PCICA. Finally, in sections V and VI we present the simulations conclusions to the chapter, respectively.

5.2 System Model

We consider the uplink channel of a multicell system with L cells, each containing a single base station (BS) with n_r antennas serving τ single antenna users, through k channel uses. All users attempt to communicate with their BS at the same time/frequency. Due to interference from neighboring cells, the signals received by a one of the BS can be expressed as,

$$\mathbf{Y} = \sqrt{\gamma}\mathbf{HBX} + \mathbf{Z}, \tag{5.1}$$

in which $\mathbf{Y} \in \mathbb{R}^{n_r \times k}$. Matrix $\mathbf{B} \in \mathbb{R}^{L\tau \times L\tau}$ is a diagonal matrix with elements β_l for $l = 1, \ldots, L\tau$ representing both path-loss and shadowing effects. The largest element β_1 is normalized as one in this chapter. Matrices $\mathbf{H} \in \mathbb{R}^{n_r \times L\tau}$ and $\mathbf{Z} \in \mathbb{R}^{n_r \times k}$ represent the fast fading gains and the independent identically distributed (i.i.d.) zero-mean circularly symmetric Gaussian noise, respectively. The channel gains are assumed to be slow block fading and remain constant through k channel uses. The matrix $\mathbf{X} \in \mathbb{R}^{l \times k}$ represents the transmitted signals. Each row of $\mathbf{X}$ is drawn from $\mathscr{X}^k$ in which $\mathscr{X} = \{2i - 1 - M : i = 1 \ldots M\}$. The constant γ is the average received signal to noise rate (SNR).

We use real values for the representation of constellations and baseband instead of complex numbers to reduce complexity of the algorithm. It is important to note that any equation of the form $\mathbf{Y}_c = \mathbf{H}_c\mathbf{X}_c + \mathbf{Z}_c$ over the complex field can be expressed through its real-valued representation as,

$$\begin{bmatrix} R_e(\mathbf{Y}_c) \\ I_m(\mathbf{Y}_c) \end{bmatrix} = \begin{bmatrix} R_e(\mathbf{H}_c) & -I_m(\mathbf{H}_c) \\ I_m(\mathbf{H}_c) & R_e(\mathbf{H}_c) \end{bmatrix} \begin{bmatrix} R_e(\mathbf{X}_c) \\ I_m(\mathbf{X}_c) \end{bmatrix} + \begin{bmatrix} R_e(\mathbf{Z}_c) \\ I_m(\mathbf{Z}_c) \end{bmatrix}. \tag{5.2}$$

Therefore if the increase in the matrices and vectors dimensions are taken into account, then (5.2) allows us to use complex channels and M-QAM modulations in this framework. We consider that the BS has no knowledge of the channel state information (CSI) and the algorithm estimates it from the received signals $\mathbf{Y}$. The only information available to the BS is the shape and average power of the constellations $\left(\rho = \mathbb{E}[\mathbf{x}_l^2]\right)$ and the total number of users.

Due to the inherent ambiguities of the fastICA and the shape of the constellations which are invariant under sign flips and permutations, it is impossible to recover the exact matrix $\mathbf{X}$. Our goal is to retrieve $\hat{\mathbf{X}} \approx \mathbf{TX}$ in which $\mathbf{T}$ is the product of a permutation matrix and a diagonal matrix with elements $(-1, 1)$ representing the possible permutation and sign flips of the data streams. Sign ambiguities can then be solved making use of a single pilot symbol [28], [29], or using differential modulation. Permutation ambiguities can be solved through the use of a small number of pilots [28], through some knowledge of the channel statistics [29] or in some cases it can be neglected if the source of the messages can be identified in upper layers of the transmission protocol.

5.3 FastICA for blind communication systems

5.3.1 PCA Whitening

The goal of this preprocessing algorithm is two-fold. First, it is necessary to whiten the received data, that is, remove correlation among the components and normalize the variance to unity. The second goal is to reduce the complexity of the fastICA by extracting the main components of $\mathbf{Y}$.

Given the sample covariance matrix $\mathbf{\Sigma} = \frac{1}{k}\mathbf{Y}\mathbf{Y}^T$ we calculate its singular value decomposition as, $\mathbf{\Sigma} = \mathbf{U}\mathbf{D}\mathbf{U}^T$ in which $\mathbf{U} \in \mathbb{R}^{n_r \times n_r}$ is an orthogonal matrix and $\mathbf{D} \in \mathbb{R}^{n_r \times n_r}$ is a diagonal matrix containing the singular values. If $n_r > l$, it is possible to divide $\mathbf{U}$ into a signal $\left(\mathbf{U}_s \in \mathbb{R}^{n_r \times l}\right)$ and noise $\left(\mathbf{U}_z \in \mathbb{R}^{n_r \times n_r - l}\right)$ subspaces. Similarly, we can divide the singular value matrix into a signal subspace ($\mathbf{D}_s$, with the l highest singular values) and a noise subspace ($\mathbf{D}_z$, with the $n_r - l$ lowest singular values). Therefore, we can express the

singular value decomposition as,

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_z \end{bmatrix} \begin{bmatrix} \mathbf{D}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_z \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^T \\ \mathbf{U}_z^T \end{bmatrix}. \tag{5.3}$$

The received signal $\mathbf{Y}$ is whitened and projected onto the signal subspace through the following expression,

$$\mathbf{Y}_w = \mathbf{D}_s^{-\frac{1}{2}} \mathbf{U}_s^T \mathbf{Y}. \tag{5.4}$$

The whitened signal $\mathbf{Y}_w$ is an $l \times k$ matrix, whereas the received data $\mathbf{Y}$ presented dimension $n_r \times k$. This reduction in dimensions significantly decreases the complexity of fastICA if $n_r >> l$.

5.3.2 FastICA

Given a whitened linear combinations of statistically independent non-Gaussian sources, ICA searches for an orthogonal matrix $\mathbf{W}$ such that,

$$\mathbf{X}_e = \sqrt{\rho} \mathbf{W}^T \mathbf{Y}_w, \tag{5.5}$$

in which $\hat{\mathbf{X}} \approx \mathbf{TX}$ is obtained after passing the data $\mathbf{X}_e$ through a slicer. Matrix $\mathbf{W}$ is obtained by minimizing the statistical dependency between the components of $\mathbf{X}_e$. The fastICA estimates one independent component (column of $\mathbf{W}$) at a time. Following [45] the p-th ($p = 1 \ldots, l$) orthogonal vector is obtained as follows,

1. Randomly choose a unit norm vector $\mathbf{w}_p$
2. Let $\mathbf{w}_p^+ = \frac{1}{k} \sum_{j=1}^{k} \left(\mathbf{y}_{w,j} g\left(\mathbf{w}_p^T \mathbf{y}_{w,j} \right) \right) - \frac{1}{k} \sum_{j=1}^{k} \left(\dot{g}\left(\mathbf{w}_p^T \mathbf{y}_{w,j} \right) \right) \mathbf{w}_p^T$.

   Here $g(.)$ represents a function that measures the nongaussianity of the estimated data ($\mathbf{w}_p^T \mathbf{Y}_w$). In the framework of this problem, it has been proven [46] that maximizing nongaussianity equates to minimizing dependency of the data streams. The function $\dot{g}(.)$ is the derivative of g and $\mathbf{y}_{w,j}$ the j-th column of $\mathbf{Y}_w$.
3. Let $\mathbf{w}_p^+ = \mathbf{w}_p^+ / \|\mathbf{w}_p^+\|$
4. If $\mathbf{w}_p^+$ does not converges, go back to step 2.

   Note that convergence here implies that the values of $\mathbf{w}_p^+$ and $\mathbf{w}_p$ point to the same direction, i.e, their dot product is approximately unitary.
5. $\mathbf{w}_p^+ = \mathbf{w}_p^+ - \sum_{i=1}^{p-1} \mathbf{w}_i^T \mathbf{w}_i \mathbf{w}_p^+$.

   This step performs a Gram-Schmidt like decorrelation. This is done to assure that the columns of $\mathbf{W}$ are orthogonal.

6. $\mathbf{w}_p = \mathbf{w}_p^+ / \|\mathbf{w}_p^+\|$

If $p = l$ the algorithm finishes and outputs $\mathbf{W}$. Otherwise $p = p + 1$ and returns to 1. The channel is estimated as,

$$\hat{\mathbf{G}}^\dagger = \rho^{\frac{1}{2}} \mathbf{W}^T \mathbf{D}_s^{-\frac{1}{2}} \mathbf{U}_s^T. \tag{5.6}$$

Assuming the use of a zero forcing (ZF) linear receiver, channel decoupling is performed as,

$$\mathbf{X}_e = \hat{\mathbf{G}}^\dagger \mathbf{Y}. \tag{5.7}$$

Finally, using a simple slicer that rounds real number to the nearest odd integer we obtain the data estimate,

$$\hat{\mathbf{X}} = 2\lfloor \mathbf{X}_e / 2 \rfloor - 1. \tag{5.8}$$

5.4 Phase Corrected ICA (PCICA)

5.4.1 FastICA and Constellation Rotation

Several factors interfere in the performance of fastICA in the system model assumed. A small SNR, a high condition number of $\mathbf{G}$ and a small symbol block (k) can decrease the accuracy of the demixing matrix ($\mathbf{W}$). Since $\mathbf{W}$ is orthogonal, performing the inner product $\mathbf{W}^T \mathbf{Y}_w$ equates to rotating the whitened data. If $\mathbf{W}$ is not accurate, then $\mathbf{X}_e$ results in a phase rotated noisy constellation.

PCICA consists in using a metric to evaluate if this estimated noisy constellation is rotated, replacing demixing vectors $\mathbf{w}_p$ if necessary. As fastICA performs local optimizations, it may generate different demixing vector each time it is used. We reduce phase rotation and thus achieve better performance iterating steps 1 to 6 until vectors that meet our criterion are found.

5.4.2 Mean Value Criterion

In our method, after a demixing vector is estimated in step 6 of fastICA, we decouple the data streams by calculating $\mathbf{x}_{e,p} = \rho^{1/2} \mathbf{w}_p^T \mathbf{Y}_w$. From $\mathbf{x}_{e,p}$ we select the elements that are at a unitary distance to $M - 1$, the largest symbol. These points are chosen for our analysis because they are the most susceptible to rotation issues. We name the vector containing these elements $\mathbf{x}_{tr,p}$ and proceed to calculate its sample mean value,

$$\mu_{s,p} = \frac{1}{n} \sum_{r=1}^{n} \mathbf{x}_{tr,p}. \tag{5.9}$$

If the demixing vector is accurate and the SNR is high, $\mu_{s,p}$ is supposed to approximate $\mu_{e,p}$, the statistical mean value of the random variable (RV) $X_{tr,p}$. This RV represents the decoupled symbols at unitary distance from $M-1$ considering perfectly accurate channel acquisition $\left(\hat{\mathbf{G}} = \mathbf{T}^{\dagger}\mathbf{G}\right)$. To analyze $\mu_{s,p}$, we perform a Z-test [42] with the following hypothesis,

$$H_0 : \mu_{s,p} = \mu_{e,p} \tag{5.10}$$

$$H_1 : \mu_{s,p} \neq \mu_{e,p}. \tag{5.11}$$

For a set significance value α, if H_0 is unlikely to be true we reject the demixing vector $\mathbf{w}_p$.

5.4.3 Performing the Z-Test

To perform the Z-test, we first have to obtain the probability density function (PDF) of $X_{tr,p}$. Considering perfect CSI at the receiver, from (5.7) we have the following equation,

$$\mathbf{X}_e = \mathbf{T}\mathbf{X} + \rho^{\frac{1}{2}}\mathbf{W}^T\mathbf{D}_s^{-\frac{1}{2}}\mathbf{U}_s^T\mathbf{Z}, \tag{5.12}$$

in which $\mathbf{Z}_{ef} = \rho^{\frac{1}{2}}\mathbf{W}^T\mathbf{D}_s^{-\frac{1}{2}}\mathbf{U}_s^T\mathbf{Z}$ is the effective noise. The noise value for a specific symbol of the p-th data stream at the i-th channel use is calculated as,

$$z_{ef,p,i} = \rho^{\frac{1}{2}} \sum_{l=1}^{n_r} \sum_{q=1}^{l} \left(w_{p,q} D_{s,q}^{-\frac{1}{2}} U_{s,l,q}^T \right) Z_{l,i} \tag{5.13}$$

in which $w_{a,b}$, $U_{s,a,b}^T$, $Z_{a,b}$, represents the element on the a-th row and b-th column of matrices $\mathbf{W}, \mathbf{U}_s^T, \mathbf{Z}$ respectively. Constant $D_{s,q}$ is the q-th diagonal element of $\mathbf{D}_s$. As $z_{i,j} \sim \mathcal{N}(0,1)$, $z_{ef,p,i}$ are also normal r.v with zero mean and variance expressed as,

$$\text{var}(z_{ef,p,i}) = \sigma_p^2 = \rho \sum_{l=1}^{n_r} \left(\sum_{q=1}^{l} \left(w_{p,q} D_{s,q}^{-\frac{1}{2}} U_{s,l,q}^T \right) \right)^2. \tag{5.14}$$

Therefore the elements of $\mathbf{X}_e$ are normally distributed around $\mathscr{X}^k$ with variance defined in (5.14).

The RV $X_{tr,p}$ can thus be seen as a truncated normal distribution. It is derived from a normal RV with $M-1$ mean, σ_p^2 variance that is bound at $[M-2, M]$. The bounding occurs due to the selection of elements at unitary distance from $M-1$. The PDF of X_{tr} is given as

$$f_{X_{tr,p}}(x) = \begin{cases} \dfrac{e^{-\frac{(-M+x+1)^2}{2\sigma_p^2}}}{\sqrt{2\pi}\sigma_p \text{erf}\left(\frac{1}{\sqrt{2}\sigma_p}\right)}, & M-2 < x < M \\ 0, & \text{otherwise} \end{cases} \tag{5.15}$$

From (5.15), we calculate the mean of $X_{tr,p}$ as,

$$\bar{X}_{tr,p} = \int_{-\infty}^{\infty} f_{X_{tr,p}}(x) x dx = M-1. \tag{5.16}$$

The second moment can be calculated as,

$$\mathbb{E}\left[X_{tr,p}^2\right] = \int_{-\infty}^{\infty} f_{X_{tr,p}}(x) x^2 dx = -\frac{\sqrt{\frac{2}{\pi}} e^{-\frac{1}{2\sigma_p^2}} \sigma_p}{\operatorname{erf}\left(\frac{1}{\sqrt{2}\sigma_p}\right)} + (M-1)^2 + \sigma_p^2, \tag{5.17}$$

therefore the variance is given as,

$$\sigma_{tr,p}^2 = \mathbb{E}\left[X_{tr,p}^2\right] - \bar{X}_{tr,p}^2 = \sigma_p^2 - \frac{\sqrt{\frac{2}{\pi}} e^{-\frac{1}{2\sigma_p^2}} \sigma_p}{\operatorname{erf}\left(\frac{1}{\sqrt{2}\sigma_p}\right)}. \tag{5.18}$$

Due to the central limit theorem, if the number of elements (n) in $\mathbf{x}_{tr,p}$ is large, then $\mu_{s,p}$ converges to a random variable with Gaussian distribution with mean $M-1$ and variance $\sigma_{tr,p}^2/n$. Since $\alpha = \mathbb{P}\left(\mu_{s,p} > t_h\right)$ it can be calculated as,

$$\alpha = 2\int_{t_h}^{\infty} \frac{e^{-\frac{(-M+x+1)^2}{2\sigma_{tr,p}^2}}}{\sqrt{2\pi}\sigma_{tr,p}} dx = \operatorname{erfc}\left(\frac{\sqrt{n}(t_h - M + 1)}{\sqrt{2}\sigma_{tr,p}}\right). \tag{5.19}$$

Here t_h represents the critical value. Fixing α we obtain the expression,

$$t_h = \frac{\sqrt{2}\,\operatorname{erfc}^{-1}(\alpha)\sigma_{tr,p}}{\sqrt{n}} + M - 1 \tag{5.20}$$

Finally, substituting (5.14) into (5.20) we have the critical value expressed as,

$$t_h = \sqrt{2}\operatorname{erfc}^{-1}(\alpha)\sqrt{\frac{\sigma_p\left(\sigma_p - \frac{\sqrt{\frac{2}{\pi}} e^{-\frac{1}{2\sigma_p^2}}}{\operatorname{erf}\left(\frac{1}{\sqrt{2}\sigma_p}\right)}\right)}{n}} + M - 1, \tag{5.21}$$

with σ_p defined in (5.14). We analyze $\mathbf{w}_p$ using the inequality,

$$|M - 1 - \mu_{s,p}| > t_h - M + 1. \tag{5.22}$$

If the inequality is satisfied, it means that it is unlikely that H_0 is true and therefore we reject $\mathbf{w}_p$. It is important to note that the assumption of CSI at the receiver is only used to derive the statistics necessary to perform the Z-test. PCICA only requires knowledge of the shape and average power of the constellation used by each user.

5.4.4 Phase Corrected ICA

Algorithm 2 presents the PCICA. The parameter $cmax$ represents the maximum number of rejected demixing vectors allowed. If $cmax$ is reached, the algorithm goes through all the rejected demixing vectors and keeps the one that generates the smallest $|M-1-\mu_{s,p}|$. This is performed to limit the complexity of the algorithm, as in some cases many iterations are needed to find a demixing vectors that satisfies our mean restriction. Low SNR and channel matrices with high condition number usually increase the number of rejected demixing vectors.

Data: Whitened data matrix $\mathbf{Y}_w$, significance level α, whitening matrices $\mathbf{D}_s$ and $\mathbf{U}_s^T$, average SNR ρ, maximum number of rejected vectors cmax.
Result: Demixing matrix $\mathbf{W}$.
for $p = 1 : l$ **do**
  Counter $= 0$;
  $\mu_{min} = \infty$;
  while *True* **do**
    Estimate $\mathbf{w}_p$ using fastICA;
    $\mathbf{x}_{e,p} = \sqrt{\rho}\mathbf{w}_p^T\mathbf{Y}_w$;
    Select all elements of $\mathbf{x}_{e,p}$ which are at unitary distance from $M-1$ and store them in $\mathbf{x}_{tr,p}$;
    $n = \text{length}(\mathbf{x}_{tr,p})$;
    $\mu_{s,p} = \frac{1}{n}\sum_{r=1}^{n}\mathbf{x}_{tr,p}$;
    $\sigma_p = \sqrt{\rho\sum_{l=1}^{n_r}\left(\sum_{q=1}^{l}\left(\mathbf{w}_{p,q}\mathbf{D}_{s,q}^{-\frac{1}{2}}\mathbf{U}_{s,l,q}^T\right)\right)^2}$;
    $t_h = \sqrt{\frac{2}{n}}\text{erfc}^{-1}(p)\sqrt{\sigma_p\left(\sigma_p - \frac{\sqrt{\frac{2}{\pi}}e^{-\frac{1}{2\sigma_p^2}}}{\text{erf}\left(\frac{1}{\sqrt{2}\sigma_p}\right)}\right)} + M - 1$;
    if $\left|M-1-\mu_{s,p}\right| < t_h - M + 1$ **then**
      Break;
    else
      if $|M-1-\mu_{s,p}| < \mu_{min}$ **then**
        $\mu_{min} = |M-1-\mu_{s,p}|$
        $\mathbf{w}_{min} = \mathbf{w}_p$
      end
      Counter $=$ Counter $+ 1$;
      if *Counter*$\geq$ *cmax* **then**
        $\mathbf{w}_p = \mathbf{w}_{min}$
        Break;
      end
    end
  end
end

Algorithm 2: Phase Corrected ICA

5.5 Simulations

In this section, we present the result of simulations performed with PCICA. We consider two scenarios in which the number of cells L and the number of users in each cell τ variates. We analyze the performance of the BS in the center cell, considering all users located outside of it as interferers. We also consider a $4-$QAM modulation, $n_r = 128$, $cmax = 20$, $\beta_a = 1$ $(a = 1 \ldots L)$ and $\beta_i = 0.8$ $(i = L+1 \ldots \tau L)$. High β_i were chosen to represent a scenario with high interference. As cell sizes become smaller in next-generation wireless systems, interference will become inevitable. For ICA based methods the BS decodes messages from all τL users even though only messages from the users in the same cell are of interest. For MMSE it was considered that the BS has perfect CSI of the users in its cell, while treating all interference as noise. Channel gains $h_{i,j}$ are complex Gaussian with zero mean and unit variance. The contrast function used is $g(u) = \tanh(u)$. Figure 5.1 shows the symbol error rate (SER) curves. It is possible to observe that PCICA signifi-

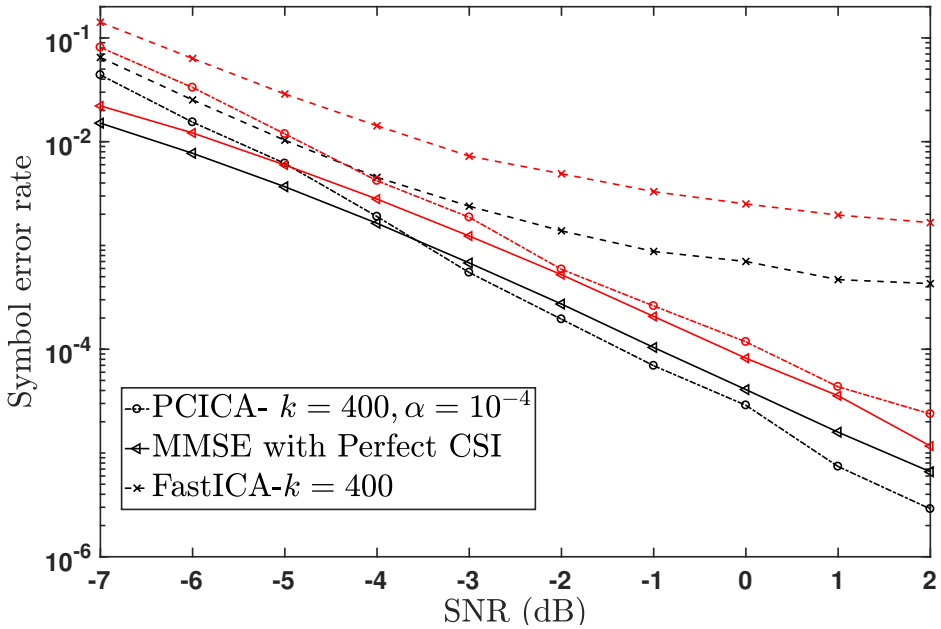

Figure 5.1 – SNR vs Symbol error rate. Scenario 1 (black curves) $L = 1, \tau = 7$. Scenario 2 (red curves) $L = 2, \tau = 4$

cantly outperforms fastICA for all considered SNR. It also outperforms MMSE at high SNR in scenario 1 and is very close in scenario 2. This occurs due to the high interference of the simulated scenarios. As we consider MMSE with only CSI of the user on its own cell, it is unable to remove extra-cell interference. For PCICA on the other hand, high interference is beneficial as it attempts to decode all data streams. If extra-cell users have high SNR, the probability that PCICA makes decoding errors of their data streams decreases. In Figure 5.2 we analyze the added complexity of PCICA compared to fastICA by showing the percentage

of rejected demixing vectors for different SNR. At high SNR, the algorithm rejects around 12% of demixing vectors for scenario 1 and around 19% for scenario 2. Thus, the PCICA consumes only a few extra iterations.

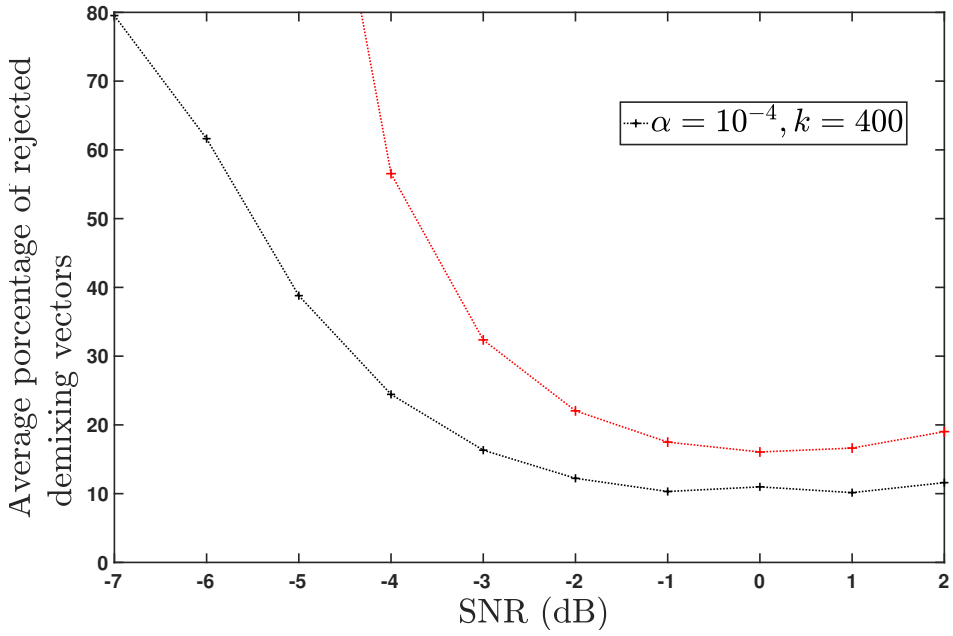

Figure 5.2 – SNR vs Average number of rejected demixing vectors. Scenario 1 (red curves) $L = 1, \tau = 7$. Scenario 2 (red curves) $L = 2, \tau = 4$.

5.6 Chapter Conclusion

In this chapter, we proposed a new algorithm to perform blind decoding. We have shown through simulations that our algorithm outperforms the traditional fastICA. For the high SNR region, we have shown that the added complexity is small.

Conclusions

This work is divided into two parts. In the first part the statistics of the sum-rate of IF and other lattice based decoders for MIMO systems is analyzed. In the second part we present contributions to blind decoders for Massive MIMO systems.

In chapter 3, we have introduced two new approximations to the achievable sum rate of the IF linear receiver in the presence of Rayleigh distributed fading. We have shown, through a series of simulations, that for the 2×2 array, over 98% of channel realizations require only one iteration of the Gauss-Lagrange algorithm to obtain λ_2. We have also shown, that our approximation for the 2×2 array, which is based on the Gauss-Lagrange algorithm, is very tight for all SNR values for both real and complex channels.

We have analytically shown that the $n \times n$ approximation attains the same degrees of freedom as the IF linear receiver and that the coefficient of variation is null for asymptotically high SNR. Through numerical simulations, we have also shown that this approximation is tight for high SNR, and its coefficient of variation is small even for limited SNR. Observing the relation of the simulations and analytical results, it is possible to state that the approximation is close to the performance of IF linear receivers regardless of channel correlation and number of antennas. From the simulations, it was also possible to observe how the IF linear receiver is more efficient and more robust to channel correlation than the zero forcing linear receiver.

In chapter 4, we have presented an approximation to the IF and CFT receivers average sum rate in the presence of channel estimation error for fixed channel realizations. We have also derived an approximation to the Ergodic IF sum rate for correlated and uncorrelated channels. Through a series of simulations, it was shown that the approximations presented are tight for all SNR values and small channel estimation error coefficient e values. Most important is the fact that even for small values of e, the system suffers a serious degradation of its sum rate.

In chapter 5, we proposed a new algorithm to perform blind decoding. The algorithm is a variation to fastICA which takes into consideration the shape of the constellation being used to iteratively select demixing vectors that lead to smaller bit error rates. We have shown through simulations that our algorithm outperforms the traditional fastICA. For the high SNR region, we have shown that the added complexity is small.

Future Work

Opportunities for future investigation are summarized below:

1. An extension to the second approximation of the IF sum-rate which is also suitable to $n \times m$ array systems.

2. In [47] the authors analyze an uplink channel where each base station (BS) is equipped with a large but finite number of antennas. It analyzes matched filer (MF) and MMSE receivers to derive how many antennas per BS are needed to achieve η% of the ultimate performance. A similar analysis could be performed to IF decoders.

3. In simulations we performed, we observed that fastICA or PCICA used jointly with IF decoders perform worse than with ZF and MMSE decoders. The reasons for this are still not well understood.

4. Instead of iteratively selecting vectors in PCICA, a new version of PCICA in which the mean value criterion is used as a restriction into the optimization problem could be researched.

References

[1] Cisco, “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update,” 2017. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

[2] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 8 2002. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6770094

[3] E. Telatar, “Capacity of Multi-antenna Gaussian Channels,” *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 11 1999. [Online]. Available: http://doi.wiley.com/10.1002/ett.4460100604

[4] B. Nazer and M. Gastpar, “Compute-and-Forward: Harnessing Interference Through Structured Codes,” *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, 10 2011. [Online]. Available: http://ieeexplore.ieee.org/document/6034734/

[5] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, “Integer-Forcing Linear Receivers,” *IEEE Transactions on Information Theory*, vol. 60, no. 12, pp. 7661–7685, 12 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6918518/

[6] O. Ordentlich and U. Erez, “Cyclic-Coded Integer-Forcing Equalization,” *IEEE Trans. on Inform. Theory*, vol. 58, no. 9, pp. 5804 – 5815, 2012.

[7] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, “Integer-Forcing Linear Receivers: A New Low-Complexity MIMO Architecture,” *2010 IEEE Vehicular Technology Conference (VTC 2010-Fall)*, pp. 1 – 5.

[8] M. Taherzadeh, A. Mobasher, and A. K. Khandani, “Communication Over MIMO Broadcast Channels Using Lattice-Basis Reduction,” *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4567–4582, 12 2007. [Online]. Available: http://ieeexplore.ieee.org/document/4385786/

[9] T. L. Marzetta, “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 11 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5595728/

[10] F. Rusek, D. Persson, Buon Kiong Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 1 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6375940/

[11] J. Zhang, X. Yuan, and Y.-J. A. Zhang, "Blind Signal Detection in Massive MIMO: Exploiting the Channel Sparsity," *IEEE Transactions on Communications*, vol. 66, no. 2, pp. 700–712, 2 2018. [Online]. Available: http://ieeexplore.ieee.org/document/8063440/

[12] K. Mawatwal, D. Sen, and R. Roy, "A Semi-Blind Channel Estimation Algorithm for Massive MIMO Systems," *IEEE Wireless Communications Letters*, vol. 6, no. 1, pp. 70–73, 2 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7752822/

[13] G. C. Ferrante, G. Geraci, T. Q. S. Quek, and M. Z. Win, "Group-Blind Detection for Uplink of Massive MIMO Systems," *IEEE Transactions on Signal Processing*, vol. 65, no. 5, pp. 1272–1286, 3 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7676246/

[14] O. Ordentlich, U. Erez, and B. Nazer, "The Approximate Sum Capacity of the Symmetric Gaussian K-User Interference Channel," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3450–3482, 6 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6787064/

[15] ——, "The compute-and-forward transform," in *2012 IEEE International Symposium on Information Theory Proceedings*. IEEE, 7 2012, pp. 3008–3012. [Online]. Available: http://ieeexplore.ieee.org/document/6284113/

[16] ——, "Successive Integer-Forcing and its Sum-Rate Optimality," in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 10 2013, pp. 282–292.

[17] A. Sakzad, J. Harshan, and E. Viterbo, "On complex LLL algorithm for integer forcing linear receivers," in *2013 Australian Communications Theory Workshop (AusCTW)*. IEEE, 1 2013, pp. 13–17. [Online]. Available: http://ieeexplore.ieee.org/document/6510037/

[18] S. Amin, J. Harshan, and V. Emanuele, "Integer-Forcing MIMO Linear Receivers Based on Lattice Reduction," *IEEE Transactions on Wireless Communication*, vol. 12, no. 10, pp. 4905 – 4915, 2013.

[19] L. Ding, K. Kansanen, Y. Wang, and J. Zhang, "Exact SMP Algorithms for Integer-Forcing Linear MIMO Receivers," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 6955 – 6966, 2015.

[20] I. El Bakoury and B. Nazer, "The impact of channel variation on integer-forcing receivers," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 6 2015, pp. 576–580. [Online]. Available: http://ieeexplore.ieee.org/document/7282520/

[21] O. Ordentlich and U. Erez, "Precoded integer-forcing universally achieves the MIMO capacity to within a constant gap," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 323 – 340, 2015.

[22] A. Sakzad and E. Viterbo, "Full Diversity Unitary Precoded Integer-Forcing," *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4316–4327, 8 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7078928/

[23] W. He, B. Nazer, and S. Shamai, "Uplink-downlink duality for integer-forcing," in *2014 IEEE International Symposium on Information Theory*. IEEE, 6 2014, pp. 2544–2548. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6875293

[24] S. M. Azimi-Abarghouyi, M. Nasiri-Kenari, B. Maham, and M. Hejazi, "Integer Forcing-and-Forward Transceiver Design for MIMO Multipair Two-Way Relaying," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 11, pp. 8865–8877, 11 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7384521/

[25] C. Wang, E. K. S. Au, R. D. Murch, W. Ho, R. S. Cheng, and V. Lau, "On the Performance of the MIMO Zero-Forcing Receiver in the Presence of Channel Estimation Error," *IEEE Transactions on Wireless Communication*, vol. 6, no. 3, 2007.

[26] K. N. Pappi, G. K. Karagiannidis, and R. Schober, "How sensitive is compute-and-forward to channel estimation errors?" in *2013 IEEE International Symposium on Information Theory*. IEEE, 7 2013, pp. 3110–3114. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6620798

[27] D. Mi, M. Dianati, L. Zhang, S. Muhaidat, and R. Tafazolli, "Massive MIMO Performance with Imperfect Channel Reciprocity and Channel Estimation Error," *IEEE Transactions on Communications*, pp. 1–1, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7867037/

[28] Y. Liu, H. Wang, W. Zhang, Q. Xu, and L. Shen, "Decoding Method Based on Complex ICA for a Multicell Massive MIMO Uplink System," *IEEE Signal*

Processing Letters, vol. 23, no. 5, pp. 648–652, 5 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7438799/

[29] L. Shen, Y.-D. Yao, H. Wang, and H. Wang, "Blind Decoding Based on Independent Component Analysis for a Massive MIMO Uplink System in Microcell Rician/Rayleigh Fading Channels," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8322–8330, 10 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7353216/

[30] T. Peken, G. Vanhoy, and T. Bose, "Blind channel estimation for massive MIMO," *Analog Integrated Circuits and Signal Processing*, vol. 91, no. 2, pp. 257–266, 5 2017. [Online]. Available: http://link.springer.com/10.1007/s10470-017-0943-1

[31] R. Zamir, B. Nazer, Y. Kochman, and I. Bistritz, *Lattice Coding for Signals and Networks*. Cambridge: Cambridge University Press, 2014. [Online]. Available: http://ebooks.cambridge.org/ref/id/CBO9781139045520

[32] Wen Zhang, Sanzheng Qiao, and Yimin Wei, "HKZ and Minkowski Reduction Algorithms for Lattice-Reduction-Aided MIMO Detection," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5963–5976, 11 2012. [Online]. Available: http://ieeexplore.ieee.org/document/6256756/

[33] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices, and groups*, 3rd ed. Springer, 1998.

[34] U. Erez and R. Zamir, "Achieving 1/2 log (1+SNR) on the AWGN channel with lattice encoding and decoding," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2293–2314, 10 2004. [Online]. Available: http://ieeexplore.ieee.org/document/1337105/

[35] O. Ordentlich, U. Erez, and B. Nazer, "The Approximate Sum Capacity of the Symmetric Gaussian K-User Interference Channel," *IEEE Trans. on Information Theory*, vol. 60, no. 6, pp. 3450 – 3482, 2014.

[36] A. S. Guerreiro, G. Fraidenraich, and S. Kumar, "Approximate Sum Rate for Integer-Forcing Receiver," *IEEE Transactions on Communications*, pp. 1–1, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7983405/

[37] R. Zamir, *Lattice Coding of Signals and Networks*, 1st ed. Cambridge, 2014.

[38] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005. [Online]. Available: https://www.cambridge.org/br/academic/subjects/engineering/wireless-communications/fundamentals-wireless-communication?format=HB

[39] E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Transactions Telecomm.*, vol. 10, no. 6, pp. 585 – 595, 1999.

[40] C. Marco, W. Moe Z., and Z. Alberto, "On the Capacity of Spatially Correlated MIMO Rayleigh-Fading Channels," *IEEE Trans. on Inform. Theory*, vol. 49, no. 10, pp. 2363 – 2371, 2003.

[41] S. Victor, "NTL: A Library for doing Number Theory-http://www.shoup.net/ntl/," 2016. [Online]. Available: http://www.shoup.net/ntl/

[42] S. M. Kay, *Fundamentals of statistical signal processing*. Prentice-Hall PTR, 1993.

[43] G. J. Székely and M. L. Rizzo, "The distance correlation t-test of independence in high dimension," *Journal of Multivariate Analysis*, vol. 117, pp. 193–213, 5 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0047259X13000262

[44] D. S. Michalopoulos, N. D. Chatzidiamantis, R. Schober, and G. K. Karagiannidis, "The Diversity Potential of Relay Selection with Practical Channel Estimation," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 481–493, 2 2013. [Online]. Available: http://ieeexplore.ieee.org/document/6409502/

[45] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications." *Neural networks : the official journal of the International Neural Network Society*, vol. 13, no. 4-5, pp. 411–30. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/10946390

[46] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent component analysis*. J. Wiley, 2001. [Online]. Available: https://www.wiley.com/en-us/Independent+Component+Analysis-p-9780471405405

[47] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?" in *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 9 2011, pp. 545–550. [Online]. Available: http://ieeexplore.ieee.org/document/6120214/

[48] A. Edelman and B. D. Sutton, "The Beta-Jacobi Matrix Model, the CS Decomposition, and Generalized Singular Value Problems," *Foundations of Computational Mathematics*, vol. 8, no. 2, pp. 259–285, 4 2008. [Online]. Available: http://link.springer.com/10.1007/s10208-006-0215-9

[49] A. M. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Arguments*. WORLD SCIENTIFIC, 10 1997. [Online]. Available: https://www.worldscientific.com/worldscibooks/10.1142/3438

Appendices

APPENDIX A

Appendix

A.1 Proof of Theorem 3.1

The matrix $\mathbf{W} = \mathbf{H}'\mathbf{H}$ is governed by the Wishart density,

$$\mathscr{P}_{\mathbf{W}}(\mathbf{W}) = C_{\beta,n} \exp\left(-\frac{\beta}{2}\mathrm{tr}\mathbf{W}\right) \det(\mathbf{W})^{\beta/2-1} H(\mathbf{W}). \tag{A.1}$$

Here $C_{\beta,n}$ is the normalization factor,

$$C_{\beta,n}^{-1} = 2^{(2-\beta)n} \pi^{\beta n(n-1)/4} \prod_{j=1}^{n} \Gamma\Big(\frac{\beta j}{2}\Big), \tag{A.2}$$

and the step function $H(\mathbf{W})$ enforces the positive-definiteness condition for the matrix $\mathbf{W}$. Now, the matrix $\mathbf{X} = \mathbf{I}_n + s\mathbf{W}$ will be distributed according to the density,

$$\begin{aligned}\mathscr{P}_{\mathbf{X}}(\mathbf{X}) = &\frac{C_{\beta,n}}{s^{\frac{\beta n^2}{2}}} \exp\left(-\frac{\beta}{2s}\mathrm{tr}(\mathbf{X}-\mathbf{I}_n)\right) \times \\ &\det(\mathbf{X}-\mathbf{I}_n)^{\frac{\beta}{2}-1} H(\mathbf{X}-\mathbf{I_n}).\end{aligned} \tag{A.3}$$

The inverse $\mathbf{Y}$ of $\mathbf{X}$ will be distributed according to,

$$\begin{aligned}\mathscr{P}_{\mathbf{Y}}(\mathbf{Y}) = &\frac{C_{\beta,n}}{s^{\beta n^2/2}} \exp\left(-\frac{\beta}{2s}\mathrm{tr}(\mathbf{Y}^{-1}-\mathbf{I}_n)\right) \times \\ &\det(\mathbf{Y}^{-1}-\mathbf{I}_n)^{\beta/2-1} \det(\mathbf{Y})^{-\beta(n-1)-2} \times \\ &H(\mathbf{Y^{-1}}-\mathbf{I_n}).\end{aligned} \tag{A.4}$$

The factor $\det(\mathbf{Y})^{-\beta(n-1)-2}$ is the Jacobian of the inversion of a real-symmetric/complex-Hermitian matrix [48]. Finally, applying Cholesky decomposition, $\mathbf{Y} = \mathbf{L}'\mathbf{L}$, we get the probability density of $\mathbf{L}$ or $\mathbf{L}'$ as presented in (3.26). The factor $2^n \prod_{j=1}^{n} l_{jj}^{\beta(n-j)+1}$ is the Jacobian associated with Cholesky decomposition, where l_{jj} represents the diagonal elements of $\mathbf{L}$ or $\mathbf{L}'$ [48], [49].

A.2 Proof of theorem 3.2

Since $l_{11}, l_{22} > 0$, we find that

$$C_{1,2} = 1/(4\pi),$$

$$\mathrm{tr}((\mathbf{L}'\mathbf{L})^{-1}-\mathbf{I}_2))=\frac{l_{11}^2+l_{12}^2+l_{22}^2-2l_{11}^2l_{22}^2}{l_{11}^2l_{22}^2},$$
$$\det((\mathbf{L}'\mathbf{L})^{-1}-\mathbf{I}_2))=\frac{1-l_{11}^2-l_{12}^2-l_{22}^2+l_{11}^2l_{22}^2}{l_{11}^2l_{22}^2},$$
$$\det(\mathbf{L}'\mathbf{L})=l_{11}^2l_{22}^2.$$

Also, since a 2×2 matrix is positive definite if its trace and determinant are positive, it follows that the positive-definiteness condition $H((\mathbf{L}'\mathbf{L})^{-1}-\mathbf{I}_n)$ requires $\mathrm{tr}(\mathbf{L}'\mathbf{L}-\mathbf{I}_2)>0$ and $\det(\mathbf{L}'\mathbf{L}-\mathbf{I}_2)>0$. These two conditions combined with $l_{11},l_{22}>0$, are equivalent to fixing the domain of the matrix elements as,

$$\begin{aligned}0<&l_{11},l_{22}<1,\\ -\sqrt{1-l_{11}^2-l_{22}^2+l_{11}^2l_{22}^2}<&l_{12}<\sqrt{1-l_{11}^2-l_{22}^2+l_{11}^2l_{22}^2}.\end{aligned}\tag{A.5}$$

The joint PDF, $p_{jointr}(l_{11},l_{12},l_{22})$, of the elements of matrix $\mathbf{L}$ for the real case is therefore expressed as,

$$p_{jointr}(l_{11},l_{12},l_{22})=\frac{1}{\pi s^2\, l_{11}^3l_{22}^4\sqrt{1-l_{11}^2-l_{12}^2-l_{22}^2+l_{11}^2l_{22}^2}}\exp\left(-\frac{l_{11}^2+l_{12}^2+l_{22}^2-2l_{11}^2l_{22}^2}{2s\, l_{11}^2l_{22}^2}\right),\tag{A.6}$$

subject to the restrictions (A.5). The cumulative density function (CDF) of τ_1 and τ_2 given in (3.12) and the joint CDF of τ_1,τ_2 for the real case are expressed respectively as,

$$F_{\tau_1}(T_1)=\int_0^1\int_{-1}^1\int_0^1 p_{jointr}(l_{11},l_{12},l_{22})H(T_1-l_{11})H(T_1-l_{22})\,\mathrm{d}l_{22}\mathrm{d}l_{12}\mathrm{d}l_{11}.\tag{A.7}$$

$$\begin{aligned}F_{\tau_2}(T_2)=&\int_0^1\int_{-1}^1\int_0^1 p_{jointr}(l_{11},l_{12},l_{22})H\left(T_2-\sqrt{l_{12}^2+l_{22}^2}\right)\times\\ &H\left(T_2-\frac{l_{11}l_{22}}{\sqrt{l_{12}^2+l_{22}^2}}\right)\mathrm{d}l_{22}\mathrm{d}l_{12}\mathrm{d}l_{11}.\end{aligned}\tag{A.8}$$

$$\begin{aligned}F_{\tau_1,\tau_2}(T_1,T_2)=&\int_0^1\int_{-1}^1\int_0^1 p_{jointr}(l_{11},l_{12},l_{22})H(T_1-l_{11})H\left(T_2-\sqrt{l_{12}^2+l_{22}^2}\right)\\ &H\left(T_2-\frac{l_{11}l_{22}}{\sqrt{l_{12}^2+l_{22}^2}}\right)H(T_1-l_{22})\,\mathrm{d}l_{22}\mathrm{d}l_{12}\mathrm{d}l_{11}.\end{aligned}\tag{A.9}$$

The CDF of $\hat{\lambda}_2$ (3.11) for the real case can be expressed as a function of the CDFs above,

$$F_{\hat{\lambda}_{2r}}(u)=F_{\tau_1}(u)+F_{\tau_2}(u)-F_{\tau_1,\tau_2}(u),\tag{A.10}$$

or directly as presented in (3.27).

A.3 Proof of Theorem 3.3

Defining the complex matrix $\mathbf{L}$ as

$$\mathbf{L} = \begin{pmatrix} l_{11} & l_{12} + i\,k_{12} \\ 0 & l_{22} \end{pmatrix}, \tag{A.11}$$

with $l_{11}, l_{22} > 0$, we find that,

$$C_{2,2} = 1/\pi,$$
$$\mathrm{tr}((\mathbf{L}'\mathbf{L})^{-1} - \mathbf{I}_2)) = \frac{l_{11}^2 + l_{12}^2 + k_{12}^2 + l_{22}^2 - 2l_{11}^2 l_{22}^2}{l_{11}^2 l_{22}^2},$$
$$\det((\mathbf{L}'\mathbf{L})^{-1} - \mathbf{I}_2)) = \frac{1 - l_{11}^2 - l_{12}^2 - k_{12}^2 - l_{22}^2 + l_{11}^2 l_{22}^2}{l_{11}^2 l_{22}^2},$$
$$\det(\mathbf{L}'\mathbf{L}) = l_{11}^2 l_{22}^2.$$

In order to force the trace and determinant to be positive to ensure positive definiteness of $\mathbf{L}$, we fix the domain of the matrix elements as,

$$\begin{aligned} &0 < l_{11}, l_{22} < 1, \\ &-\sqrt{1 - l_{11}^2 - l_{22}^2 + l_{11}^2 l_{22}^2} < l_{12} < \sqrt{1 - l_{11}^2 - l_{22}^2 + l_{11}^2 l_{22}^2}, \\ &-\sqrt{1 - l_{11}^2 - l_{12}^2 - l_{22}^2 + l_{11}^2 l_{22}^2} < k_{12} < \sqrt{1 - l_{11}^2 - l_{12}^2 - l_{22}^2 + l_{11}^2 l_{22}^2}. \end{aligned} \tag{A.12}$$

The joint PDF, $p_{jointc}(l_{11}, l_{12}, l_{22}, k_{2,1})$, of the elements of matrix $\mathbf{L}$ for the complex case is therefore expressed as,

$$p_{jointc}(l_{11}, l_{12}, l_{22}, k_{2,1}) = \frac{1}{\pi s^4 l_{11}^5 l_{22}^7} \exp\left(-\frac{l_{11}^2 + l_{12}^2 + k_{12}^2 + l_{22}^2 - 2l_{11}^2 l_{22}^2}{s\, l_{11}^2 l_{22}^2}\right), \tag{A.13}$$

subject to restrictions (A.12).

Having p_{jointc}, it is possible to derive the CDF of $\hat{\lambda}_2$ for the complex case using the same reasoning as in the previous subsection and therefore it can be expressed as (3.28).